Pruitt, W.E. (1981): General one-sided laws of the iterated logarithm. Ann. Probab. **9**, No. 1, 1–48. Zbl. 462.60030

Révész, P. (1967): The Laws of Large Numbers. Akadémiai Kiadó, Budapest and NewYork, Academic Press, 1968. Zbl. 203.50403

Rosalsky, A. (1980): On the converse to the iterated logarithm law. Sankhya **A42**, No. 1–2, 103–108. Zbl. 486.60031

Rozovskii, L.V. (1975): Asymptotic expansions in the central limit theorem. Teor. Veroyatn. Primen. **20**, No. 4, 810–820. Engl. transl.: Theory Probabl. Appl. **20**, Mno. 4, 794–804. Zbl. 347.60020

Rozovskii, L.V. (1978a): On a lower bound for the remainder in the central limit theorem. Mat Zametki **24**, No. 3, 403–410. Engl. transl.: Math. Notes **24**, No. 3–4, 715–719. Zbl. 396.60026

Rozovskii, L.V. (1978b): On the exactness of an estimate of the remainder term in the central limit theorem. Teor. Veroyatn. Primen. **24**, No. 4, 744–761. Engl. transl.: Theory Probabl. Appl. **23**, No. 4, 712–730. Zbl. 388.60026

Sakhanenko, A.I. (1984): Rate of convergence in an invariance principle for non-identically distributed variables with exponential moments. Tru. Inst. Mat. **3**, 4–49. Engl. transl.: Advances in Pribability Theory, Transl. for Math. Engl. 2–73 (1986). Zbl. 541.60024

Sakhanenko, A.I. (1985): Estimates in an invariance principle, Tru. Inst. Mat. **5**, 27–44 (in Russian). Zbl. 585.60044

Stout, W.F. (1974): Almost Sure Convergence, Academic Press, New York. Zbl. 321.60022

Strassen, V. (1964): An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie Verw. Geb. **3**, No. 3, 211–226. Zbl. 132.12903

Strassen, V. (1966): A converse to the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie Verw. Geb. **4**, 265–268. Zbl. 141.16501

Tomkins, R.J. (1980): Limit theorems without moment hypotheses for sums of independent random variables. Ann. Probab. **8**, No. 2, 314–324. Zbl. 432.60034

Tomkins, R.J. (1991): Refinements of Kolmogorov's law of the iterated logarithm. Stat. Probab. Lett. **14**, No. 4, 321–325 (1992), Zbl. 756.60033

Weiss, M. (1959): On the law of the iterated logarithm. J. Math. Mech. **8**, No. 1, 121–132. Zbl. 091.14206

Zolotarev, V.M. (1986): Modern Theory of Summation of Independent Random Variables. Nauka, Moscow (in Russian). Engl. transl.: VSP, Utrecht (1997). Zbl. 649.60016

II. The Accuracy of Gaussian Approximation in Banach Spaces

V. Bentkus, F. Götze, V. Paulauskas and A. Račkauskas

Contents

Chapter 0
Introduction and Notation

Let B be a real separable Banach space with norm $||\cdot|| = ||\cdot||_B$. Suppose that X, X_1, X_2, $\ldots \in B$ are independent and identically distributed (i.i.d.) random elements (r.e.'s) taking values in B. Furthermore, assume that $EX = 0$ and that there exists a zero-mean Gaussian r.e. $Y \in B$ such that the covariances of X and Y coincide. Define

$$S_n = S_n(X) = n^{-1/2}(X_1 + \cdots + X_n).$$

Let $f : B \to \mathbb{R}$ (or more generally $f : B \to F$, where F is a real Banach space) denote a function such that the expectations $EF(S_n)$, $Ef(Y)$ are well defined. The central limit theorem (CLT) in this context requires

$$\lim_{n\to\infty} Ef(S_n) = Ef(Y) \tag{0.1}$$

for all functions $f \in \mathcal{F}$, where $\mathcal{F}$ is some class of functions $f : B \to \mathbb{R}$. In the classical definition of the CLT, $\mathcal{F}$ coincides with the class $C_b(B)$ of bounded continuous functions $f : B \to \mathbb{R}$. In recent approaches $\mathcal{F}$ can differ from $C_b(B)$ and the expectation is understood not only as the traditional Lebesgue and Bochner integrals but also as other integrals.

Denote by ρ a semi–metric (i.e., $\rho(x, y) = 0$ does not necessary imply $x = y$) on a class of probability measures on B. Another formulation of the CLT is given by

$$\lim_{n\to\infty} \rho(\mathcal{L}(S_n), \mathcal{L}(Y)) = 0, \tag{0.2}$$

where $\mathcal{L}(X)$ denotes the distribution of the r.e. X. If ρ is the Prokhorov metric, then (0.2) is equivalent to the classical CLT. In what follows we shall consider the CLT in the classical sense.

The aim of the paper is to give a review of results and methods used to estimate the rate of convergence in the CLT in infinite-dimensional spaces. A review of results concerning asymptotic expansions is also presented. It is appropriate to remark that the earlier books by Sazonov (1981), Koroliuk and Borovskikh (1984) and Paulauskas and Račkauskas (1989) contain related reviews. In this paper we restrict ourselves to the case of i.i.d. summands and sums having a Gaussian limit. We do not try to give the most general or accurate formulations (for this we refer to original papers). Rather, we shall stress the differences between the finite-dimensional case and infinite-dimensional case, emphasizing new phenomena arising in connection with the large class of geometrical structures in Banach spaces. Thus the paper contains mainly sketches of proofs and many technical arguments are skipped. Our intention was also to reflect an increasing number of applications of Banach space results to certain asymptotic problems of mathematical statistics.

We shall use the following abbreviations:

$$\begin{aligned} \varepsilon &= \varepsilon(n) = n^{-1/2}, \ n = 1, 2, \ldots, \\ g(\varepsilon) &= g(\varepsilon, f) = Ef(S_n), \\ g(0) &= g(0, f) = Ef(Y), \\ h(\varepsilon) &= h(\varepsilon, \rho) = \rho(\mathcal{L}(S_n), \mathcal{L}(Y)). \end{aligned}$$

Then (0.1) means that the function $g(\varepsilon)$ of the discrete argument $\varepsilon = n^{-1/2}$, $n = 1, 2, \ldots$ is continuous at the point $\varepsilon = 0$:

$$\lim_{\varepsilon\to 0} g(\varepsilon) = g(0). \tag{0.3}$$

Similarly, (0.2) is equivalent to

$$\lim_{\varepsilon\to 0} h(\varepsilon) = 0. \tag{0.4}$$

A stronger assertion than (0.3) is to assume that g at the point $\varepsilon = 0$ satisfies a Hölder condition (say, with an exponent $\alpha > 0$ and a constant $C < \infty$)

$$|g(\varepsilon) - g(0)| \le C\varepsilon^\alpha, \tag{0.5}$$

(similarly,

$$h(\varepsilon) \le C\varepsilon^\alpha \tag{0.6}$$

in the case of (0.4)). The inequalities (0.5), (0.6) are usually called estimates of the convergence rate in the CLT. Clearly, one can rewrite (0.5) and (0.6) as

$$|Ef(S_n) - Ef(Y)| \le Cn^{-\alpha/2}, \tag{0.7}$$

$$\rho(\mathcal{L}(S_n), \mathcal{L}(Y)) \le Cn^{-\alpha/2}. \tag{0.8}$$

More precise approximations of $g(\varepsilon)$ than (0.5), (0.6) are given by the asymptotic expansion

$$g(\varepsilon) = g(0) + a_1\varepsilon + \cdots + a_k\varepsilon^k + R \tag{0.9}$$

with coefficients $a_1, \ldots, a_k$ and the remainder term $R = R_k(\varepsilon)$ such that $R_k(\varepsilon) = o(\varepsilon^k)$ as $\varepsilon \to 0$. Clearly, one can rewrite (0.9) as

$$Ef(S_n) = Ef(Y) + a_1 n^{-1/2} + \cdots + a_k n^{-k/2} + R. \tag{0.10}$$

Relations like (0.10) are usually called asymptotic expansions in the CLT.

In the first chapter of the paper we give a review of methods and results concerning estimate (0.7) when f is the indicator function of a set $A \subset B$: $f(x) = \chi_A(x)$. Then (0.7) yields an estimate for probabilities

$$|P\{S_n \in A\} - P\{Y \in A\}| \le Cn^{-\alpha/2}. \tag{0.11}$$

When the space B has finite dimension one can prove (0.11) for relatively large classes of sets A, for instance for the class of all convex sets (see, e.g., Bhattacharya and Rao (1976), Sazonov (1981)). The situation changes dramatically in the infinite–dimensional case. In this case there does not exist a uniform estimate of type (0.11) even for such relatively small classes as the class of all balls or the class of all halfspaces (see Sazonov (1981)). Therefore we have to consider special sets A. Let $F : B \to \mathbb{R}$ be a functional and let $A = A_r(F) = \{x \in B : F(x) < r\}$. Then (0.11) yields

$$|P\{F(S_n) < r\} - P\{F(Y) < r\}| \le Cn^{-\alpha/2}. \qquad (0.12)$$

It turns out that in the infinite-dimensional case the estimate (0.12) depends strongly on the smoothness properties of F. The Fourier method and the method of integration by parts discussed in §§1.1, 1.3 are adapted to the investigation of (0.12) for smooth functions F, while the Lindeberg method and the method of finite-dimensional approximation discussed in §§1.2, 1.4 are useful for the case of non–smooth functions F. The last section of the chapter discusses the results concerning the convergence rate in the CLT, estimated by means of the Prokhorov and bounded Lipschitz (BL) metrics.

The second chapter is devoted to asymptotic expansions. In §§2.1 and 2.2 we consider asymptotic expansions for the expectation $Ef(S_n)$ with $f : B \to \mathbb{R}$ a sufficiently smooth function or f a function having isolated points of non–differentiability. The expectation $E||S_n||^p$, $p > 0$, with a sufficiently smooth norm–function, is a typical example. The results discussed in these sections are obtained with no explicit condition like the classical one of Cramér. In §2.3 asymptotic expansions for $P\{F(S_n) < r\}$ are reviewed. In §2.4 we consider asymptotic expansions for the density function $(d/dr)P\{F(S_n) < r\}$.

In the third chapter we present examples applying the methods and results described in the previous chapters. We emphasize that our aim is to give illustrations of applications only. Here we distinguish the following situations:

i) cases where a limit theorem for a statistical test can be interpreted as a particular case of a general result in Banach spaces. The ω^2–statistic is a typical example of this kind (see §3.1). Every such statistic can be represented as the norm of the sum of i.i.d.r.e's in L_2 space;

ii) cases where a problem can be reduced to known facts from probability theory in Banach spaces. L–statistics (i.e. linear combinations of rank statistics) are a typical example (see §3.2). Such statistics can be represented as a sum of i.i.d. $\mathbb{R}$–valued r.v's with an additional remainder term. The remainder term can be majorized by the norm of a sum of Banach-space-valued r.e's;

iii) cases where methods and ideas from probability theory in Banach space can be used in statistical problems. The typical examples here are the so–called U–statistics;

iv) cases where it is necessary to use a combination of the previous approaches. The estimation of the convergence rate for empirical processes is an example (see §3.3 and §3.4).

If not stated otherwise, we suppose that all random elements and random variables under consideration are independent. We refer to Vakhaniya et al. (1987) and Araujo and Giné (1980) for the general information on probability theory and the CLT in Banach spaces, respectively.

Let us introduce the notation we shall use throughout.

- B, F — real separable Banach spaces with norm $||\cdot|| = ||\cdot||_B$, $||\cdot|| = ||\cdot||_F$.
- B^* — the dual Banach space of B consisting of all linear continuous functionals with the standard *sup* norm.
- H — a real separable Hilbert space. We denote the scalar product by $(\cdot,\cdot)$ and the norm by $||\cdot|| = ||\cdot||_H$.
- $\mathbb{R}$ — the real line.
- $\mathbb{R}^k$ — the k–dimensional Euclidean space; $(x, y) = x_1y_1 + \cdots + x_ky_k$ denotes the scalar product.
- ℓ_p, c_0 — the classical Banach spaces of sequences.
- $L_p(S, S, \nu)$, L_p, $C[0,1]$, $D[0,1]$ — the classical spaces of functions.
- a.s. — almost surely.
- i.i.d. — independent and identically distributed.
- r.e. — random element.
- B–r.e. — r.e. with values in B.
- r.v. — random variable ($\mathbb{R}$–r.e.)
- c.f. — characteristic function
- $X, X_1, X_2, \ldots \in B$ — a sequence of i.i.d. B–r.e.'s such that $EX = 0$
- $Y, Y_1, Y_2, \ldots \in B$ — a sequence of i.i.d. Gaussian B–r.e.'s such that $EY = 0$ and the covariances of X and Y coincide, i.e. $Ef(X)g(X) = Ef(Y)g(Y)$ for all $f, g \in B^*$.
- $S_n = S_n(X) = n^{-1/2}(X_1 + \cdots + X_n)$.
- $\mu = \mathcal{L}(X)$ — the distribution of X.
- $\nu = \mathcal{L}(Y)$ — the distribution of Y.
- $cov\, X$ — the covariance operator of X. By definition $cov\, X : B^* \to B$, $cov\, X(f) = Ef(X)X$ for all $f \in B^*$.
- $\beta_s = E||X||^s$, $s \in \mathbb{R}$ — the s–th moment of X.
- $\nu_s = \nu_s(\mathcal{L}(X), \mathcal{L}(Y)) = \int_B ||x||^s |\mathcal{L}(X) - \mathcal{L}(Y)|(dx)$ — the s–th pseudomoment.
- $|\mu - \nu|$ — the variation of the signed measure $\mu - \nu$.
- $\chi_A(t)$ — the indicator function of a set A.
- $f^{(s)}(x)$ — the s–th Fréchet derivative of the function f at the point x.
- $f^{(s)}(x)h_1 \ldots h_s = f^{(s)}(x)[h_1, \ldots, h_s]$ — the value of the derivative $f^{(s)}(x)$ as an s–linear continuous form at the point $(h_1, \ldots, h_s)$.
- $f^{(s)}(x)h^s = f^{(s)}(x)[h^s] = f^{(s)}(x)h \ldots h$.
- $||f^{(s)}(x)|| = \sup\{||f^{(s)}(x)h^s|| : ||h|| \le 1\}$ — the norm of the s–linear form $f^{(s)}(x)$.
- $C^s = C^s(B; F)$ — the space of all s–times continuously Fréchet differentiable functions $f : B \to F$.
- $C_b^s = C_b^s(B; F)$ — the space of all bounded functions in C^s having bounded derivatives.

Chapter 1
Rates of Convergence

§1.1. The Fourier Method

The Fourier method has a long tradition in mathematics and particularly in probability. It is especially effective when used to investigate the distribution of sums of independent r.v.'s. In number theory it was employed by Gauss to obtain quadratic reprocity laws and to study the representation of integers by sums of squares or higher powers (the Waring problem); for the latter, use was made of the so-called Hardy–Littlewood circle method in analytic number theory (Hardy and Littlewood (1920)). In two papers, Lyapunov (1900, 1901) adopted the Fourier method to prove the CLT on the real line under the so-called Lyapunov moment conditions. This method relies on inequalities of the following type between the difference of probability measures and its Fourier transform (called the characteristic function or c.f. for short). Let μ be a probability measure and ν a measure of finite variation on $\mathbb{R}$. If ν is absolutely continuous with respect to Lebesgue measure and its density p is bounded: $m = \sup_{x\in\mathbb{R}} |p(x)| < \infty$, then for each $T > 0$ we have

$$\sup_{x\in\mathbb{R}} |\mu((-\infty, x)) - \nu((-\infty, x))| \le C_1 \int_{-T}^{T} \frac{|\hat{\mu}(t) - \hat{\nu}(t)|}{|t|}\, dt + C_2 \frac{m}{T}, \tag{1.1}$$

where C_1 and C_2 are absolute constants and the c.f.

$$\hat{\nu}(t) = \int_{\mathbb{R}} \exp(itx)\nu(dx).$$

This is the so-called Berry–Esseen lemma (Berry (1941), Esseen (1942)). It is used to prove convergence rates and higher order approximations to a measure μ by a measure ν with smooth density by means of their characteristic functions. Here T is chosen so that T^{-1} has a desired error size. The advantage of the Fourier method in Lyapunov's approach is apparent from the proof of the classical Berry–Esseen estimate

$$\sup_{x\in\mathbb{R}} |P\{S_n < x\} - P\{Y < x\}| = O(n^{-1/2}), \tag{1.2}$$

where $S_n = n^{-1/2}(X_1 + \cdots + X_n)$ and $X, X_1, X_2, \ldots$ are i.i.d. real r.v.'s such that

$$EX = 0, \quad \beta_3 := E|X|^3 < \infty$$

and Y denotes a Gaussian r.v. such that $EY = 0$ and $\beta_2 := EX^2 = EY^2 > 0$. Setting in (1.1) $\mu = \mu_n = \mathcal{L}(S_n)$, $\nu = \mathcal{L}(Y)$ and $T = A\sqrt{n}$ for some $A = A(\mathcal{L}(Y)) > 0$ to be chosen later on, we can reduce (1.2) to the estimate

$$\int_{-T}^{T} \frac{|\hat{\mu}_n(t) - \hat{\nu}(t)|}{|t|}\, dt = O(n^{-1/2}). \tag{1.3}$$

Usually the estimation of the difference $\hat{\mu}_n(t) - \hat{\nu}(t)$ is based on different arguments for "small" and "large" values of $|t| \le T$. Clearly

$$\hat{\mu}_n(t) = E\exp(itS_n) = \Big(E\exp(itn^{-1/2}X)\Big)^n.$$

Expanding the exponential in a Taylor series we obtain

$$\Big|E\exp(itn^{-1/2}X)\Big| \le |1 - \beta_2 t^2/2n| + \beta_3|t|^3/6n^{3/2} \le \exp(-\beta_2 t^2/3n),$$

$$|\hat{\mu}_n(t)| \le \exp(-\beta_2 t^2/3) \tag{1.4}$$

for $|t| \le An^{1/2}$ with $A = \min(\beta_2/\beta_3,\ \sqrt{2\beta_2})$. Relation (1.4) and the equality $\hat{\nu}(t) = \exp(-\beta_2 t^2/2)$ allow one to show that on every interval $n^\varepsilon \le |t| \le A\sqrt{n}$ the estimate (1.3) holds for each fixed $\varepsilon > 0$. For "small" values of $|t| \le n^\varepsilon$ one should additionally take into account the coincidence of the moments of X and Y up to second order, that is

$$\frac{d^s}{dt^s}\left(\hat{\mu}_n(t) - \hat{\nu}(t)\right)\Big|_{t=0} = 0, \quad s = 0, 1, 2.$$

For example, one can expand

$$\ln \hat{\mu}_n(t) = n \ln E\exp(itn^{-1/2}X)$$

in a Taylor series in a neighborhood of zero,

$$n \ln E\exp(itn^{-1/2}X) = -\beta_2 t^2/2 + O(\beta_3|t|^3 n^{-1/2}),$$

and obtain

$$\frac{|\hat{\mu}_n(t) - \hat{\nu}(t)|}{|t|} = O\Big(t^2\exp(-t^2\beta_2/2)n^{-1/2}\Big).$$

This guarantees the necessary estimate $O(n^{-1/2})$ of the integral (1.3) on the interval $|t| \le n^\varepsilon$. We refer the reader to Lyapunov (1900), Esseen (1945), Ibragimov and Linnik (1971) and Petrov (1975) for more information on the Fourier method in the one-dimensional case.

Similar argumentation is possible in the finite-dimensional case. Extensions of the Berry–Esseen Lemma are obtainable there (see, e.g., Bhattacharya and Rao (1976), Bhattacharya (1977), Sweeting (1977) and Sazonov (1981)). The c.f.

$$\hat{\mu}_n(t) = E\exp(i(t, S_n)) = \Big(E\exp(in^{-1/2}(t, X))\Big)^n,$$

where $(x, y) = x_1y_1 + \cdots + x_ky_k$, $x, y \in \mathbb{R}^k$, is the standard scalar product and r.e. $X \in \mathbb{R}^k$, still has the multiplicative structure and, generally speaking, one can repeat the one-dimensional arguments with certain, sometimes very complicated, technical changes.

Now let X, X_1, X_2, ... denote r.e.'s taking values in a measurable space $(\mathcal{X}, \mathcal{B})$. Instead of sums one can study more general statistics

$$T = t_n(X_1, \ldots, X_n), \tag{1.5}$$

which are symmetric in $X_1, \ldots, X_n$ and such that the influence of each X_j is asymptotically negligible. An example of (1.5) is provided by U-statistics (see Götze (1979), Bickel, van Zwet and Götze (1986), Götze (1987)). Another example is

$$T = F(S_n),$$

where the summands X, X_1, X_2, ... of S_n take values in a separable Banach space B and are such that $EX = 0$, $E||X||^s < \infty$, $s \geq 2$. Here $F : B \to \mathbb{R}$ denotes a function that has enough Fréchet derivatives. Therefore T is no longer a sum of i.i.d. r.v.'s and the c.f. of T does not have a multiplicative structure. Thus a new technique for analyzing the c.f. of T has to be developed to still be able to use the Berry–Esseen lemma. Let us start with the simplest infinite-dimensional case where B denotes a Hilbert space H and $F(x) = ||x||^2$. The following extremely useful symmetrization inequality allows one to reduce the analysis of $E\exp(it||S_n||^2)$ to that of a product of certain characteristic functions (see Lemma 1.8 for generalizations).

The symmetrization $\bar{U}$ of a r.e. U is defined as $\bar{U} = U_1 - U_2$, where U_1 and U_2 are independent copies of r.e. U.

Lemma 1.1 (Götze (1979)). *For arbitrary independent r.e.'s U, V, $W \in H$ and $t \in \mathbb{R}$, the inequality*

$$|E\exp(it||U + V + W||^2)|^4 \leq E\exp(2it(\bar{U}, \bar{V}))$$

holds.

Proof. Note that for any real-valued function $f(u, v)$,

$$|E\exp(itf(U, V))|^2 \leq E\exp(it\Delta_2(V_1 - V_2)f(U, V_2)), \tag{1.6}$$

where V_1, V_2 are independent copies of V and the difference operator $\Delta_2(h)f(u, v) = f(u, v+h) - f(u, v)$ acts on the second argument of f. Indeed, applying Hölder's inequality and using the fact that V, V_1, V_2 are i.i.d., we have

$$\begin{aligned} &|E\exp(itf(U, V))|^2 \leq E|E(\exp(itf(U, V))|U)|^2 \\ &= E\exp(itf(U, V_1))\exp(-itf(U, V_2)), \end{aligned}$$

which coincides with (1.6). Applying (1.6) twice with $f(U, V) = ||U+V+W||^2$ and using in this case that

$$\Delta_1(h_1)\Delta_2(h_2)f(u, v) = 2(h_1, h_2),$$

where the difference operator Δ_1 acts on the first argument of f, we complete the proof.

In order to illustrate how the symmetrization inequality works, we sketch the proof of the following theorem.

Theorem 1.2. *Let $X, Y \in H$. Suppose that X is a bounded r.e., $P\{||X|| \leq M\} = 1$, and X is not concentrated in a finite-dimensional subspace of H. Then*

$$\Delta_n := \sup_{x>0} |P\{||S_n||^2 < x\} - P\{||Y||^2 < x\}| = O(n^{-1/2}).$$

In the proof of the theorem we assume that X is not concentrated in a subspace of H, otherwise we can replace H by this subspace. Thus, there exists an orthonormal basis $\{e_k, k \in \mathbb{N}\}$ of H such that

$$Y = \sum_{k=1}^{\infty} \sigma_k \eta_k e_k, \quad \sum_{k=1}^{\infty} \sigma_k^2 < \infty,$$

where η, η_1, η_2, ... $\in \mathbb{R}$ are i.i.d. standard normal r.v.'s, $E\eta = 0$, $E\eta^2 = 1$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$.

Lemma 1.3. *For all $t \in \mathbb{R}$ and $l = 1, 2, \ldots$,*

$$|E\exp(it||Y||^2)| \leq (1 + 4t^2\sigma_l^4)^{-l/4}. \tag{1.7}$$

Moreover,

$$|E\exp(it(Y_1, Y_2))| \leq (1 + t^2\sigma_l^4)^{-l/2}. \tag{1.8}$$

Proof. It is easy to verify that

$$E\exp(it\eta^2) = 1/\sqrt{1 - 2it}.$$

Therefore

$$E\exp(it||Y||^2) = \prod_{k=1}^{\infty}(1 - 2it\sigma_k^2)^{-1/2}$$

and this clearly implies (1.7). Furthermore, since

$$E\exp(it(x, Y)) = \exp(-(Dx, x)/2),$$

where $D = cov\, Y$, we have

$$E\exp(it(Y_1, Y_2)) = E\exp(-t^2(DY, Y)/2) \leq (E\exp(-t^2\sigma_l^4\eta^2))^l$$

which obviously implies (1.8).

Lemma 1.4. *If $P\{||X|| \leq M\} = 1$, then for all $p > 0$ and m, l such that $m + l \leq n$ the following inequality holds:*

$$E||n^{-1/2}(X_1 + \cdots + X_m + Y_1 + \cdots + Y_l)||^p \leq C(p)M^p.$$

Proof. It is sufficient to prove that

$$E\|n^{-1/2}(Y_1+\cdots+Y_l)\|^p \le C(p)M^p, \tag{1.9}$$

$$E\|n^{-1/2}(X_1+\cdots+X_m)\|^p \le C(p)M^p. \tag{1.10}$$

The inequality (1.9) follows from $\mathcal{L}(Y_1+\cdots+Y_l)=\mathcal{L}(\sqrt{l}Y)$ and

$$E\|Y\|^p \le C(p)(E\|Y\|^2)^{p/2} = C(p)(E\|X\|^2)^{p/2} \le C(p)M^p.$$

To prove (1.10) we apply the Zygmund–Marcinkiewicz inequality (see e.g. Vakhaniya et al. (1987))

$$E\|Z_1+\cdots+Z_m\|^p \le C(p)E(\|Z_1\|^2+\cdots+\|Z_m\|^2)^{p/2}$$

which holds for arbitrary i.i.d. r.e.'s $Z_1,\ldots,Z_m \in H$ with zero mean. Therefore (1.10) reduces to the obvious inequality $n^{-p/2}E(\|X_1\|^2+\cdots+\|X_m\|^2)^{p/2} \le M^p$.

Lemma 1.5. *Let $P\{\|X\|\le M\}=1$. Define $\bar U := n^{-1/2}(\bar X_1+\cdots+\bar X_m)$, $m\le n$. Suppose that the r.e.'s $Z,\bar U \in H$ are independent. Then for all $s\in\mathbb{R}$ and $L>0$ such that*

$$|s|LM \le \sqrt{n}, \tag{1.11}$$

we have

$$0 \le E\exp(is(Z,\bar U)) \le P\{\|Z\|>L\} + E\exp\left(-\frac{s^2 m}{2n}(DZ,Z)\right), \tag{1.12}$$

where $D=\operatorname{cov}X=\operatorname{cov}Y$.

Proof. Since the r.e. $\bar U$ is the symmetrization of $U:=n^{-1/2}(X_1+\cdots+X_m)$, the inequality on the left-hand side of (1.12) holds. To prove the inequality on the right-hand side of (1.12), it suffices to show that for each (non-random) $z\in H$ we have

$$E\exp(is(z,\bar U)) \le \chi\{\|z\|>L\} + \exp\left(-\frac{s^2 m}{2n}(Dz,z)\right). \tag{1.13}$$

If $\|z\|>L$, then (1.13) is obvious. Therefore we assume that $\|z\|\le L$. Then (1.13) reduces to

$$E\exp(is(z,\bar U)) \le \exp\left(-\frac{s^2 m}{2n}(Dz,z)\right). \tag{1.14}$$

We have

$$E\exp(is(z,\bar U)) = \left[E\cos\left(\frac{s}{\sqrt{n}}(s,\bar X)\right)\right]^m \tag{1.15}$$

since $\bar X,\bar X_1,\bar X_2,\ldots$ are i.i.d. and $\bar X$ is symmetric. Put $x=s(z,\bar X)/\sqrt{n}$. Then (1.11) together with $\|z\|<L$, $\|\bar X\|\le 2M$ ensure that $|x|\le 2$. Therefore in

(1.15) we can apply the obvious inequality $\cos x \le 1-x^2/4$ valid for $|x|\le 2$. Noting that $E(z,\bar X)^2 = 2(Dz,z)$ we have

$$E\exp(is(z,\bar U)) \le \left[1-\frac{s^2}{2n}(Dz,z)\right]^m \le \exp\left(-\frac{s^2 m}{2n}(Dz,z)\right),$$

which completes the proof.

Proof of Theorem 1.2. The distribution function $r\to P\{\|Y\|^2<r\}$ has a bounded density provided $\sigma_3>0$. Indeed, it follows from Lemma 1.3 that the c.f.

$$|E\exp(it\|Y\|^2)| \le \min(1;\sigma_3^{-3}|t|^{-3/2})$$

is integrable. This allows one to apply the Berry–Esseen lemma with $T=\sqrt{n}$. Therefore the estimate $\Delta_n = O(n^{-1/2})$ will result from the following four bounds:

$$I_1 := \int_{n^\varepsilon<|t|<\sqrt{n}} \frac{1}{|t|}|g(t)|\,dt = O(n^{-1/2});$$
$$I_2 := \int_{n^\varepsilon<|t|<\sqrt{n}} \frac{1}{|t|}|f_n(t)|\,dt = O(n^{-1/2});$$
$$I_3 := \int_{|t|\le 1} \frac{1}{|t|}|f_n(t)-g(t)|\,dt = O(n^{-1/2});$$
$$I_4 := \int_{1\le|t|\le n^\varepsilon} \frac{1}{|t|}|f_n(t)-g(t)|\,dt = O(n^{-1/2}),$$

where $0<\varepsilon\le\frac12$ is a number to be chosen later and

$$g(t) := E\exp(it\|Y\|^2), \quad f_n(t) := E\exp(it\|S_n\|^2).$$

Let us estimate I_1 first. It follows from Lemma 1.3 that

$$|g(t)| \le 1/(\sigma_l^l|t|^{l/2}).$$

Therefore $I_1 \le 2/(l\sigma_l^l n^{\varepsilon l/2}) = O(n^{-1/2})$ if $\varepsilon l\ge 1$. But such an $l=l(\varepsilon)$ exists since $\sigma_l>0$ for all l.

Next we shall estimate I_2. It is sufficient to show for $1\le|t|\le\sqrt{n}$ and a sufficiently large fixed constant $A>0$ that

$$|f_n(t)| = O(|t|^{-A}+n^{-A}). \tag{1.16}$$

Write $S_n = U+V+W$, where

$$U=n^{-1/2}(X_1+\cdots+X_m), \quad V=n^{-1/2}(X_{m+1}+\cdots+X_{m+k}), \quad W=S_n-U-V$$

and $k+m\le n$. Then by the symmetrization inequality (see Lemma 1.1), (1.16) follows from

$$E\exp(it(\bar V,\bar U)) = O(|t|^{-A}+n^{-A}). \tag{1.17}$$

Using Lemma 1.5 we obtain for all $L > 0$

$$E\exp(it(\bar V,\bar U)) \le P\{||\bar V|| > L\} + E\exp\left(-\frac{t^2 m}{2n}(D\bar V,\bar V)\right), \tag{1.18}$$

provided $|t|LM \le \sqrt{n}$. Choose (noting that $1 \le |t| \le \sqrt{n}$)

$$L = 1/(M|t|^{1/4}), \quad k \sim n/(2|t|), \quad m = n-k \sim n(1-1/(2|t|)).$$

Using Chebyshev's inequality and Lemma 1.4, we have

$$P\{||\bar V|| > L\} = P\{||\bar S_k|| > L\sqrt{n/k}\} \le C(M,A)(k/(nL^2))^{2A} = O(|t|^{-A}). \tag{1.19}$$

Similarly,

$$E\exp\left(-\frac{t^2 m}{2n}(D\bar V,\bar V)\right) \le E\exp\left(-\frac{|t|}{8}(D\bar S_k,\bar S_k)\right)$$
$$\le P\{(D\bar S_k,\bar S_k) \le 8/\sqrt{|t|}\} + \exp(-\sqrt{|t|}). \tag{1.20}$$

Estimates (1.18)–(1.20) reduce (1.17) to

$$P\{(D\bar S_k,\bar S_k) \le 8/\sqrt{|t|}\} = O(|t|^{-A} + n^{-A}), \tag{1.21}$$

where $k \sim n/(2|t|)$, and $1 \le |t| \le \sqrt{n}$. But (1.21) is a consequence of the following concentration inequality (note that $k \ge \sqrt{n}/2$):

$$P\{(D\bar S_k,\bar S_k) < \varepsilon^2\} = O(\varepsilon^l + k^{-l/2}), \tag{1.22}$$

valid for all $\varepsilon > 0$ and $l = 1, 2, \ldots$. Such a concentration inequality for balls in Hilbert space can be reduced to the analogous statement in $\mathbb{R}^l$ via $||x||^2 \ge \sum_{s=1}^l (x,e_s)^2$. In $\mathbb{R}^l$ concentration inequalities of this kind have been proved by Esseen (1968), Paulauskas (1973), Götze (1979), etc. For the infinite-dimensional case, see, e.g., Siegel (1981), Bentkus (1985b), etc. The proof of (1.22) in $\mathbb{R}^l$ can be obtained as follows. For $s \ge 0$ we have

$$P\{||S_n||^2 \le \varepsilon^2\} \le \exp(\varepsilon^2 s^2/2)E\exp\left(-\frac{s^2}{2}||S_n||^2\right).$$

If we denote by $Z \in \mathbb{R}^l$ a r.e. with the standard normal distribution, then

$$E\exp\left(-\frac{s^2}{2}||S_n||^2\right) = E\exp(is(Z,S_n)).$$

Now one can apply Lemma 1.5.

Let us estimate I_3. Put $G(x) = \exp(it||x||^2)$. Then $G \in C^\infty$ and

$$G'''(x)h^3 = -8t^2 G(x)[it(x,h)^3 + (x,h)(h,h)]. \tag{1.23}$$

The estimate $I_3 = O(n^{-1/2})$ clearly follows from

$$|EG(S_n) - EG(Y)| = O(|t|n^{-1/2}). \tag{1.24}$$

We shall prove (1.24) using Lindeberg's method (see the beginning of § 1.2). Obviously

$$|EG(S_n) - EG(Y)| \le J_1 + \cdots + J_n,$$

where

$$J_k = |EG(W_k + n^{-1/2}X) - EG(W_k + n^{-1/2}Y)|,$$
$$W_k = n^{-1/2}(X_1 + \cdots + X_{k-1} + Y_{k+1} + \cdots + Y_n).$$

To estimate J_k, we use the Taylor formula

$$G(x+h) = G(x) + G'(x)h + \frac{1}{2}G''(x)h^2 + \frac{1}{2}E(1-\tau)^2 G'''(x+\tau h)h^3$$

with $x = W_k$, $h = n^{-1/2}X$ and $h = n^{-1/2}Y$, respectively. Here the r.v. τ is uniformly distributed on [0, 1]. The terms containing derivatives of G up to the second order vanish since the means and covariances of X and Y coincide and we have

$$J_k \le n^{-3/2}(J'_k + J''_k),$$

where

$$J'_k = |E(1-\tau)^2 G'''(W_k + \tau n^{-1/2}X)X^3|, \tag{1.25}$$
$$J''_k = |E(1-\tau)^2 G'''(W_k + \tau n^{-1/2}Y)Y^3|.$$

It follows from (1.23) for $|t| \le 1$ that

$$|G'''(x)h^3| = O(|t|\cdot||h||^3(1+||x||^3)).$$

Therefore

$$J'_k = O\left(|t|E||X||^3(1+||W_k+\tau n^{-1/2}X||^3)\right) = O(|t|)$$

since $||X|| \le M$ and, according to Lemma 1.4, we have $E||W_k||^3 = O(1)$. Similarly, $J''_k = O(|t|)$ since all of the moments of the Gaussian r.e. Y are finite. Therefore $J_k = O(n^{-3/2}|t|)$, which yields (1.24).

It remains to estimate I_4. Here we shall combine the methods employed in estimating I_2 and I_3. Repeating the estimation of I_3, we arrive at, for example, the integral J'_k (see (1.25)) and it suffices to show for $|t| \ge 1$ and each $A > 0$ that

$$J'_k = O(|t|^3(|t|^{-A} + n^{-A})). \tag{1.26}$$

It is clear from the explicit formula (1.23) for G''' that

$$G'''(W_k + \tau n^{-1/2}X)X^3 = EP(t,X,W_k,\tau n^{-1/2}X)\exp(it||W_k + n^{-1/2}\tau X||^2),$$

where $P(\cdot,\cdot,\cdot,\cdot)$ is a polynomial not exceeding third degree in each argument. Applying the triangle inequality, we can reduce the estimation of J'_k to that of a certain sum of quantities like $|t|^3\gamma$, where

$$\gamma := |E(W_k, X)^3 \exp(it||W_k + \tau n^{-1/2}X||^2)|,$$

and so (1.26) will follow from

$$\gamma = O(|t|^{-A} + n^{-A}). \tag{1.27}$$

Let us split $W_k = T_1 + T_2 + T_3 + T_4$ into four sums in such a way, that each sum T_1, T_2, T_3, T_4 has approximately the same number (equivalent to $n/4$) of terms X_j or Y_j. Then

$$(W_k, X)^3 = \sum_{1 \le l_1, l_2, l_3 \le 4} (T_{l_1}, X)(T_{l_2}, X)(T_{l_3}, X)$$

and the estimation of γ is reduced to estimating

$$\gamma_1 := |E(T_{l_1}, X)(T_{l_2}, X)(T_{l_3}, X) \exp(it||W_k + \tau n^{-1/2}X||^2)|.$$

But among T_{l_1}, T_{l_2}, T_{l_3} (with fixed l_1, l_2, l_3) at least one of T_1, T_2, T_3, T_4 is absent, say T_1. Therefore

$$\gamma_1 \le E||X||^3 \cdot ||T_{l_1}|| \cdot ||T_{l_2}|| \cdot ||T_{l_3}||\gamma_2,$$

where

$$\gamma_2 = |E_{T_1} \exp(it||T_1 + W||^2)|, \quad W = W_k + \tau n^{-1/2}X - T_1,$$

and E_{T_1} denotes the expectation taken only with respect to the r.e. T_1. We can estimate the moments of the r.e.'s T_{l_1}, T_{l_2}, T_{l_3} by Lemma 1.4. To estimate γ_2, we repeat the arguments of the estimation of $E\exp(it||S_n||^2)$ used in the proof of the relation $I_2 = O(n^{-1/2})$. Here we replace S_n by T_1 and use the fact that T_1 contains at least $n/4$ summands, and that T_1 and W are independent. In this way we arrive at $\gamma_2 = O(|t|^{-A} + n^{-A})$, which implies (1.27). This completes the proof of the theorem.

Write

$$\Delta_n(a, r) = |P\{||S_n - a||^2 < r\} - P\{||Y - a||^2 < r\}|,$$
$$\Delta_n(a) = \sup_{r \ge 0} \Delta_n(a, r).$$

Estimates of $\Delta_n(a, r)$ in Hilbert space were obtained in a number of papers by various methods. For $\beta_3 = E||X||^3 < \infty$ and fixed $a \in H$, the bounds on $\Delta_n(a)$ were improved from logarithmic order in Kandelaki (1965) to $O(n^{-1/6})$ by Paulauskas (1976b) and to $O(n^{-1/4})$ by Yurinskii (see Sazonov (1981)). Assuming additionally that the coordinates of r.e. X are independent, Nagaev and Chebotarev (1978) showed that $\Delta_n(0) = O(n^{-1/2})$. To prove the same bound, Borovskikh and Račkauskas (1979) and Račkauskas (1981) needed only seven coordinates independent of the rest. The essential step is due to Götze (1979) who introduced the symmetrization inequality (Lemma 1.1) for investigations of $\Delta_n(a, r)$ by the Fourier method. Götze (1979) proved the bound $\Delta_n(a) = O(n^{-1/2})$ for a fixed $a \in H$ for $\beta_6 < \infty$ and the bound $\Delta_n(0) = O(n^{-1+\varepsilon})$, $\varepsilon > 0$, for $\beta_8 < \infty$. Yurinskii (1982) introduced in addition a truncation technique and obtained the following result.

Theorem 1.6. *There exists a constant $C = C(\mathcal{L}(Y)) > 0$ such that for each $a \in H$ the following inequality holds:*

$$\Delta_n(a) \le C(1 + ||a||^3)\beta_3 n^{-1/2}. \tag{1.28}$$

The Fourier method was refined and various improvements and generalizations were obtained by Zalesskii (1982, 1985), Nagaev (1983, 1985, 1988, 1989a), Bentkus (1984d, e), Sazonov and Zalesskii (1985), Nagaev and Chebotarev (1986), Senatov (1986), Sazonov, Ul'yanov and Zalesskii (1987a, b, 1988, 1989a), Aliev (1989) and Sazonov and Ul'yanov (1990, 1991). It is interesting to note than in infinite-dimensional Hilbert space, one can obtain better rates than $O(n^{-1/2})$ without conditions like the classical Cramér condition for the c.f. Namely, if $E||X||^{2+2\delta} < \infty$, $0 \le \delta < 1$, then $\Delta_n(a) = o(n^{-\delta})$ if $a = 0$ or X is symmetric (see §2.3 for details). This goes back to a result in $\mathbb{R}^k$ due to Esseen (1945) and has been proved by Zalesskii (1982) when $a = 0$ and by Bentkus (1984e) in the general case. However the following problem remains open: *is it true that for infinite-dimensional r.e. $X \in H$ the condition $E||X||^4 < \infty$ implies $\Delta_n(0) = O(n^{-1})$?*

The structure of the constant C in (1.28) has been investigated by various authors. The dependence of the estimate on the eigenvalues σ_i^2 of $cov\, Y$ was studied in particular. Here things have progressed from thirteen eigenvalues in Yurinskii (1982) to seven in Nagaev (1983), Sazonov, Ul'yanov and Zalesskii (1987) and to six in Senatov (1989b). Recently Nagaev (1989b) and Zalesskii, Sazonov and Ul'yanov (1988) gave the estimate

$$\Delta_n(a) \le \frac{C_1\beta_3}{\sigma_1 \cdots \sigma_6}(\sigma^3 + ||a||^3)n^{-1/2}, \tag{1.29}$$

where $\sigma^2 = E||X||^2$, whereas Senatov (1989a) gave the result

$$\Delta_n(a, r) \le \frac{C_1\beta_3}{\sigma_1 \cdots \sigma_6}(r^3 + \beta_3 n^{-1/2})n^{-1/2}. \tag{1.30}$$

In both estimates C_1 is an absolute constant. The proof of (1.29) it is now 1999–2000 appeared in Sazonov, Ul'yanov and Zalesskii (1989b).

Lower bounds of the convergence rate in Hilbert space were found by Senatov (1985a, b, 1986), Barsov (1987), Aliev (1987) and Bloznelis (1989). Senatov (1986) proved the necessity of the first six eigenvalues of $cov\, Y$ in (1.29) and (1.30). Barsov has shown that $\Delta_n(0) = O(n^{-(s-2)/2})$ for some $s > 2$ implies $E||X||^{s-\varepsilon} < \infty$ for each $\varepsilon > 0$, $\varepsilon \le s$.

A truncation lemma adapted for nonuniform estimates for $\Delta_n(a, r)$ was introduced by Sazonov and Zalesskii (1985). Nonuniform estimates of $\Delta_n(a, r)$ were obtained by Sazonov and Zalesskii (1985), Bentkus and Zalesskii (1985) and Sazonov, Ul'yanov and Zalesskii (1987b, 1988). The following bound provides an example of a nonuniform estimate

$$\Delta_n(a, r) = O\left(\frac{1 + ||a||^3}{(1 + \rho)^s\sqrt{n}}\right),$$

where $E||X||^s < \infty$ for some $s \geq 3$ and $\rho = |\sqrt{r} - ||a|||$ denotes the distance between 0 and the boundary of the ball $\{x \in H : ||x - a|| < \sqrt{r}\}$.

The case of non-identically distributed summands was considered in Bentkus (1984e), Ul'yanov (1987) and Borisov (1989). For results concerning the local CLT in H see §2.4.

We end the review of results obtained in Hilbert space with the following one on large deviations.

Theorem 1.7 (Yurinskii (1988)). *Suppose that*

$$E\exp(c||X||) < \infty$$

for some constant $c > 0$. *Then there exist constants* $A_i = A_i(\mathcal{L}(X)) > 0$ *such that*

$$P\{||S_n|| > r\} = P\{||Y|| > r\}I(r, n, \mathcal{L}(X))(1 + \theta A_1 r n^{-1/2})$$

for $A_2 \leq r \leq A_3 n^{1/2}$. *Here* $|\theta| \leq 1$ *and* $I(r, n, \mathcal{L}(X))$ *denotes a certain analog of the Cramér series in large deviations theorems. Furthermore,*

$$P\{||S_n|| > r\} = P\{||Y|| > r\}(1 + \theta A_4 r^3 n^{-1/2})$$

if $A_2 \leq r \leq A_5 n^{1/6}$.

For details concerning the construction of $I(r, n, \mathcal{L}(X))$, see Yurinskii (1988). In the case of Hilbert space this improves results of Osipov (1978a, b), Bentkus (1986c), Bentkus and Račkauskas (1990), Zalesskii (1989) and Račkauskas (1988).

Consider now Banach spaces. Put

$$\Delta_{n,F}(r) = |P\{F(S_n) < r\} - P\{F(Y) < r\}|,$$

where X and Y assume values in the Banach space B and a function $F : B \to \mathbb{R}$. As in the Hilbert space case the estimation of $\Delta_{n,F}$ strongly depends on the symmetrization inequalities. An example of such an inequality is provided by the following lemma.

Lemma 1.8. *If a function* $f(u_1, \ldots, u_k)$ *of arguments* $u_1, \ldots, u_k \in B$ *is real-valued and the r.e.'s* $U_1, \ldots, U_k \in B$ *are independent, then*

$$|E\exp(itf(U_1, \ldots, U_k))|^{2^k}$$
$$\leq E\exp(i\Delta_1(U_1' - U_1)\cdots\Delta_k(U_k' - U_k)f(U_1, \ldots, U_k)),$$

where U_s' *denotes an independent copy of* U_s *and the difference operator*

$$\Delta_s(h)f(\ldots, u_s, \ldots) = f(\ldots, u_s + h, \ldots) - f(\ldots, u_s, \ldots)$$

acts on the s-th argument of f.

In particular, for all independent $U_1, \ldots, U_{k+1} \in B$ *and every polynomial* $\pi(x) = \pi_k(x) + \cdots + \pi_0(x)$, $x \in B$, *where* $\pi_s(x) = \pi_s(x, \ldots, x)$, $s = 0, \ldots, k$, *denote symmetric continuous s-linear forms on B, the following holds:*

$$|E\exp(i\pi(U_1 + \cdots + U_{k+1}))|^{2^k} \leq E\exp(ik!\pi_k(\bar{U}_1, \ldots, \bar{U}_k)), \tag{1.31}$$

where $\bar{U}$ *denotes the symmetrization of* U.

Lemma 1.8 is an easy generalization of Lemma 1.1 if one uses an estimate of the type (1.6) k times. Inequality (1.31) was given in a somewhat different form by Weyl (1916) for polynomials $\pi(x)$ where x has a uniform distribution on the discrete set $\{1, \ldots, N\}$. Here it was necessary to split x in the form $x = x_1 + \cdots + x_k + x_{k+1}$, where $x_1, \ldots, x_k$ are uniformly distributed and x_{k+1} is a function of these variables, so as to yield a uniformly distributed sum. The generality of Weyl's inequality made possible numerous applications in analytic number theory (see, e.g., Schmidt (1984)). For polynomials in one variable, there are sharper estimates due to Vinogradov (1934). These seem not easily extendable to general probability distributions. The inequality (1.31) for $k = 2$ was derived for the probability context independently in Götze (1979). The immediate extension to $k > 2$ was used by Yurinskii (1983) to prove the following result.

Theorem 1.9 (Yurinskii (1983)). *Let* $B = \ell_p$, $p = 2, 4, \ldots$. *Then*

$$|P\{||S_n - a|| < r\} - P\{||Y - a|| < r\}| \leq C\beta_3\sigma^{-3}(1 + ||a/\sigma||^{3p-3})n^{-1/2},$$

where $\sigma^2 = E||X||^2$ *and* $C = C(\mathcal{L}(Y))$ *is a constant.*

To formulate the next result we need the following conditions.

Differentiability condition (D_5). There exists $p \geq 0$ such that

$$\sup_{x \in B} P(1 + ||x||)^{-p}||F^{(s)}(x)|| < \infty, \quad s = 0, \ldots, 5.$$

Variance condition (V). For sufficiently large $M > 0$ and for each fixed c

$$\sup_{||a|| \leq c} \{\sigma(Y + a) < \delta\} = O(\delta^M) \text{ as } \delta \downarrow 0,$$

where $\sigma^2(x) = E(F'(x)Y)^2$.

Theorem 1.10 (Götze (1983)). *If* $E||X||^3 < \infty$ *and* (D_5) *and* (V) *are satisfied, then*

$$\sup_{r \in \mathbb{R}} |P\{F(S_n) < r\} - P\{F(Y) < r\}| = O(n^{-1/2}). \tag{1.32}$$

Actually Götze (1983) proved a slightly more exact result. Under certain natural conditions Zalesskii (1985) made the estimate (1.32) nonuniform in the Hilbert space $B = H$. The proof of Theorem 1.9 uses the Fourier method and exploits a slightly more general symmetrization inequality as Lemma 1.8. Götze (1986) introduced the method of integration by parts (see § 1.3) which allows one to prove more precise results than Theorem 1.10. Nevertheless, generalizations and extensions of the Fourier method, used to prove Theorems like 1.9 and 1.10, are useful for the construction of asymptotic expansions (see § 2.3).

Remark 1.11. It seems that verification of variance condition (V) is not simple in general. For example, it is still unknown whether (V) (or some other suitable condition) is satisfied for each infinite-dimensional r.e. $Y \in B = L_p$, $1 < p < \infty$, with $F(x) = ||x||$ or $F(x) = ||x||^s$ with an appropriate $s > 0$.

§1.2. The Lindeberg Method

Lindeberg's proof of the central limit theorem, which appeared in 1920 and 1922, is very simple and can be easily extended to investigate the rate of convergence even for B-r.e.'s. A similar remark applies to Trotter's proof given in 1959. Actually Trotter's proof differs from Lindeberg's only in terminology, but the method is presented in an intuitively more understandable manner. Besides the original papers of Lindeberg (1920, 1922) and Trotter (1959), we mention the books of Billingsley (1968), Thomasian (1969) and Feller (1971).

In this section we shall concentrate our attention on some extensions of Lindeberg's method that have been used to investigate the rate of convergence in the CLT in Banach spaces. Recall that X, Y are B-r.e.'s such that $EX = EY = 0$, $cov\, X = cov\, Y$ and Y is Gaussian. Furthermore, $S_n = n^{-1/2}(X_1 + \cdots + X_n)$, where $X, X_1, X_2, \ldots$ are i.i.d.

Briefly, Lindeberg's approach is as follows. Suppose we have to estimate the difference of the measures $\mu_n = \mathcal{L}(S_n)$ and $\nu = \mathcal{L}(Y)$ on a certain measurable set $A \subset B$, i.e., we have to estimate the quantity

$$\mu_n(A) - \nu(A) = \int_B \chi_A(x)(\mu_n - \nu)(dx).$$

The first step is to replace the discontinuous indicator function χ_A by a sufficiently smooth function say $g = g_{A,\varepsilon}$, which coincides with χ_A everywhere with the exception of an ε-neighborhood $(\partial A)_\varepsilon = \{x \in B : \inf_{y \in \partial A} ||x - y|| < \varepsilon\}$ of the boundary ∂A of the set A. After this substitution (usually called the smoothing lemma) two terms are to be estimated: the integral $I = \int_B g(x)(\mu_n - \nu)(dx)$ and the quantity $\nu((\partial A)_\varepsilon)$. The estimation of I is based on the Taylor formula and the following identity:

$$\begin{aligned}\mu_n - \nu &= \mathcal{L}(S_n(X)) - \mathcal{L}(S_n(Y)) \\ &= \sum_{k=1}^{n} \left[\mathcal{L}(W_{n,k} + n^{-1/2}X_k) - \mathcal{L}(W_{n,k} + n^{-1/2}Y_k)\right], \qquad (2.1)\end{aligned}$$

where $W_{n,k} = n^{-1/2}\left(\sum_{i=1}^{k-1} X_i + \sum_{i=k+1}^{n} Y_i\right)$. If g can be chosen to be three-times Fréchet differentiable such that $\sup_{x \in B} ||g'''(x)|| \le C\varepsilon^{-3}$, we can expand each $g(W_{n,k} + n^{-1/2}X_k)$ and $g(W_{n,k} + n^{-1/2}Y_k)$ around $W_{n,k}$ and obtain

$$\begin{aligned}g(W_{n,k} + n^{-1/2}X_k) &= g(W_{n,k}) + g'(W_{n,k})[X_k]n^{-1/2} + \frac{1}{2}g''(W_{n,k})[X_k^2]n^{-1} \\ &\quad + \frac{1}{6}g'''(W_{n,k} + \theta X_k n^{-1/2})[X_k^3]n^{-3/2},\end{aligned}$$

where $|\theta| \le 1$. Due to the equality of the means and covariances of X_k and Y_k (see Lemma 2.2), the difference $Eg(W_{n,k} + n^{-1/2}X_k) - Eg(W_{n,k} + n^{-1/2}Y_k)$ will contain only terms involving third derivatives and we easily arrive at the bound

$$I \le Cn^{-1/2}\varepsilon^{-3}\nu_3,$$

where

$$\nu_3 = \int_B ||x||^3 |\mathcal{L}(X) - \mathcal{L}(Y)|(dx).$$

The quantity $\nu((\partial A)_\varepsilon)$ is usually of order $C\varepsilon$. Therefore, if $\varepsilon^4 = n^{-1/2}\nu_3$, we obtain the bound $C\nu_3^{1/4}n^{-1/8}$ for the difference $\mu_n(A) - \nu(A)$. Such a rate was found for the first time using Lindeberg's method in Hilbert space by Kuelbs and Kurtz (1974). It is clear, however, that in order to derive better bounds than $O(n^{-1/8})$, one needs some additional arguments.

Firstly, there are several ways to construct a smooth approximation to an indicator function such that the smoothing error is as small as possible. In finite-dimensional space, one can use the convolution of the indicator function with a Gaussian distribution. This type of smoothing leads to the convolution method, which yields Berry–Esseen-type convergence rate results. This method was applied explicitly for the first time by Bergström (1944). We refer the reader to Sazonov's book (1981) for more details on the convolution method. Bentkus (1986a) constructed directly the approximation of an indicator function by a once-differentiable function whose derivative satisfies Hölder's condition with a constant independent of the dimension and exploited the smoothing properties of the Gaussian terms in the identity (2.1). This resulted in a better dependence of the remainder term on the dimension.

However, neither the arguments of the convolution method nor those used by Bentkus have a satisfactory straightforward extension to the infinite-dimensional case. This is due to the limited smoothness properties of Gaussian measures and to the fact that there is no analogue of Lebesgue measure in infinite-dimensional Banach spaces. Efforts to provide such an extension to Hilbert space were made by Osipov and Rotar' (1985). They considered the rate of convergence on balls and obtained Berry–Esseen-type bounds with a logarithmic factor. Optimal bounds had already been found by the methods of characteristic functions (see the previous section). Nevertheless the arguments used by Osipov and Rotar' prove to be useful in the case of dependent r.e.'s (see, for example, Račkauskas (1990) for the martingale case).

If g is a smoothed indicator function such that $g(x) = 1$ if $x \in A$ and $g(x) = 0$ if $x \in B - A_\varepsilon$, then one may use the simple fact that the derivatives of g vanish outside an ε-neighborhood of the boundary ∂A. This idea has been exploited in several papers using in addition either iteration or induction arguments. It yields bounds of at most order $O(n^{-1/6})$ by assuming a finite third moment. Despite the great difference as compared to the finite-dimensional Berry–Esseen-type results, such orders are in general unimprovable (see Theorem 2.6 below). Iteration arguments appeared in Paulauskas (1976b) and under certain assumptions, he gave bounds on balls of order $O(n^{-1/6})$. Refinements and generalizations were given by Butzer et al. (1979), Ul'yanov (1981), Bernotas (1980) and Bentkus and Račkauskas (1982), (1983). Bentkus and Račkauskas (1982) applied different notions of smoothness. This allowed them to prove rates for sets with very unsmooth boundaries such as balls in the spaces $C[0,1]$, c_0, etc. These papers mostly considered the convergence rate on balls under the assumption that for $a \in B$ the function $r \to (d/dr)P\{\|Y+a\| < r\}$ is bounded and admits certain estimates. Such an assumption can be verified in Hilbert space and some other "good" Banach spaces but it may fail in general Banach spaces (especially in $C[0,1]$ and c_0). Račkauskas proved (see Paulauskas and Račkauskas (1989)) that $O(n^{-1/6})$ remains valid under the natural assumption that the density $r \to (d/dr)P\{\|Y\| < r\}$ is bounded and the third moment is finite.

We shall consider the estimation of the quantity

$$\Delta_{n,q} := \sup_{r\geq 0} |P\{q(S_n(X)) < r\} - P\{q(Y) < r\}|$$

in detail, where $q : B \to \mathbb{R}$ is a contuous semi-norm. Without loss of generality, we may assume that $q(x) \leq \|x\|$ for all $x \in B$. First we introduce some conditions needed to formulate results obtained using Lindeberg's method with induction.

Smoothness condition (A_3). For each $r \geq 0$ and $\varepsilon > 0$ there exists a function $g_{r,\varepsilon} : B \to \mathbb{R}$ such that

(a) for all $x \in B$,

$$\chi(q(x) < r) \leq g_{r,\varepsilon}(x) \leq \chi(q(x) < r+\varepsilon);$$

(b) the function $g_{r,\varepsilon}$ is three times continuously Frechét differentiable and there is a constant $C > 0$ such that for all $r \geq 0$, $\varepsilon > 0$ and $i = 1, 2, 3$,

$$\sup_{x\in B} \|g^{(i)}_{r,\varepsilon}(x)\| \leq C\varepsilon^{-i}.$$

Density condition (D). There exists a constant $C = C(q, \mathcal{L}(Y)) > 0$ such that for all $\varepsilon > 0$

$$\sup_{r\geq 0} P\{r-\varepsilon \leq q(Y) \leq r+\varepsilon\} \leq C\varepsilon.$$

Actually condition (D) is equivalent to the existence of a bounded density of the distribution function $P\{q(Y) < r\}$.

Theorem 2.1. *Suppose that conditions (A_3), (D) hold and that the Gaussian r.e. Y is infinite-dimensional. Then there exists a constant $C = C(q, \mathcal{L}(Y)) > 0$ such that*

$$\Delta_{n,q} \leq C\nu_3^{1/3} n^{-1/6}.$$

Proof. Define for $k = 1, \ldots, n$ the quantity

$$\Delta_{n,k,q} = \sup_{r\geq 0} |P\{q(U_{k,n}) < r\} - P\{q(Y) < r\}|,$$

where $U_{k,n} = n^{-1/2}(\sum_{i=1}^k X_i + \sum_{i=k+1}^n Y_i)$, with $\sum_{i\in\emptyset} \equiv 0$. We now proceed to prove inductively the bound

$$\delta_n := \max_{1\leq k\leq n} \Delta_{n,k,q} \leq C_0 \nu_3^{1/3} n^{-1/6},$$

which of course yields the desired result since $\Delta_{n,q} \leq \delta_n$.

Since Y is infinite-dimensional, we have $\delta_1 \leq C_0\nu_3^{1/3}$ (see, e.g., Theorem 5.1.11 in Paulauskas and Račkauskas (1989)). So let $n > 1$ and suppose that according to the induction assumption

$$\delta_{n-1} \leq C_0 \nu_3^{1/3} (n-1)^{-1/6}. \tag{2.2}$$

Let $\varepsilon > 0$ denote a parameter which will be specified at the end of the proof. If $r \leq 2\varepsilon$, we have $\Delta_{n,k,q}(r) \leq \Delta_{n,k,q}(2\varepsilon) + C\varepsilon$, where

$$\Delta_{n,k,q}(r) := |P\{q(U_{n,k}) < r\} - P\{q(Y) < r\}|.$$

Therefore it is sufficient to estimate $\Delta_{n,k,q}(r)$ for $r \geq 2\varepsilon$ only. Define $g_1(x) = g_{r,\varepsilon}(x)$, $g_2(x) = g_{r-\varepsilon,\varepsilon}(x)$, where $g_{r,\varepsilon}$ is the same function as in condition (A_3). Define

$$G_{n,k} = \mathcal{L}(U_{n,k}), \quad P_{n,k} = \mathcal{L}(W_{n,k}), \quad H = \mathcal{L}(X) - \mathcal{L}(Y).$$

One easily shows the following smoothing inequality (see Kuelbs and Kurtz (1974)):

$$\Delta_{n,k,q}(r) \leq \max_{i=1,2} \left| \int_B g_i(x)(G_{n,k} - G)(dx) \right| + P\{r-\varepsilon \leq q(Y) \leq r+\varepsilon\}. \tag{2.3}$$

The last probability does not exceed $C\varepsilon$ due to condition (D). To estimate the integrals in (2.3) we proceed as in Lindeberg's method. Using (2.1), we obtain

$$\left| \int_B g_i(x)(G_{n,k} - G)(dx) \right| \leq \sum_{j=1}^k \left| \int_B \int_B g_i(y + xn^{-1/2}) P_{n,j}(dy) H(dx) \right|.$$

We now expand $g_i(y + xn^{-1/2})$ in a Taylor series in a neighborhood of y obtaining

$$g_i(y+xn^{-1/2}) = g_i(y) + g_i'(y)[x]n^{-1/2} + \frac{1}{2}n^{-1}g_i''(y)[x^2] + \frac{1}{2}n^{-3/2}\int_0^1 (1-\theta)^2 g_i'''(y+\theta n^{-1/2}x)[x^3]\,d\theta. \quad (2.4)$$

Note that $\int_B g_i'(y)[x]H(dx) = 0$ since $EX = EY = 0$. Moreover,

$$\int_B g_i''(y)[x^2]H(dx) = 0$$

since the covariances of X and Y coincide. This is not trivial because $cov\, X = cov\, Y$ implies $Ef(X)g(X) = Ef(Y)g(Y)$ for all $f, g \in B^*$ only. The second derivative $g_i''(y)$, which is continuous bilinear form on B, in general may not be generated by linear functionals. But we have the following result due to Borisov (1989).

Lemma 2.2. *Let B be a separable Banach space. Let r.e.'s $X, Y \in B$, $E||X||^2 < \infty$ and Y be Gaussian. If $EX = EY = 0$ and $cov\, X = cov\, Y$, then for each continuous bilinear form T on B the following holds:*

$$ET(X, X) = ET(Y, Y).$$

We continue with the proof by integrating both sides of (2.4) with respect to $P_{n,j}(dy)$ and $H(dx)$. We obtain

$$I_j := \int_B \int_B g_i(y + n^{-1/2}x)P_{n,j}(dy)H(dx) = \frac{1}{2}n^{-3/2}\int_0^1 (1-\theta)^2 \times \int_B\int_B g_i'''(y+\theta n^{-1/2}x)[x^3]P_{n,j}(dy)H(dx)d\theta. \quad (2.5)$$

Split the integral with respect to $H(dx)$ into two parts over $\{||x|| \le \varepsilon\sqrt{n}\}$ and $\{||x|| > \varepsilon\sqrt{n}\}$. Denote the first integral by $I_{j,1}$ and the second one by $I_{j,2}$. In order to estimate $I_{j,1}$ we use the following bound on the third derivative of the function g_i, which follows from the property (a) of condition (A_3) and the assumption $q(x) \le ||x|| \le \varepsilon n^{1/2}$:

$$|g_i'''(y+\theta n^{-1/2}x)[x^3]| \le C||x||^3\varepsilon^{-3}\chi(r-\varepsilon \le q(y+\theta n^{-1/2}x) \le r+\varepsilon) \le C||x||^3\varepsilon^{-3}\chi(r-2\varepsilon \le q(y) \le r+2\varepsilon).$$

Therefore we have

$$|I_{j,1}| \le C\varepsilon^{-3}n^{-3/2}\nu_3 P_{n,j}\{r-2\varepsilon \le q(y) \le r+2\varepsilon\}. \quad (2.6)$$

Due to condition (D),

$$P_{n,j}\{r-2\varepsilon \le q(y) \le r+2\varepsilon\} \le 2\delta_{n-1} + P\{(r-2\varepsilon)(n/n-1)^{1/2} \le q(Y) \le (r+2\varepsilon)(n/n-1)^{1/2}\} \le 2\delta_{n-1} + C\varepsilon, \quad (2.7)$$

and we arrive at the estimate

$$|I_{j,1}| \le C\varepsilon^{-3}n^{-3/2}\nu_3(\delta_{n-1}+\varepsilon). \quad (2.8)$$

For $\{||x|| > \varepsilon\sqrt{n}\}$ we have

$$\chi(r-\varepsilon \le q(y+\theta n^{-1/2}x) \le r+\varepsilon) \le \chi(r - 2||x||n^{-1/2} \le q(y) \le r + 2||x||n^{-1/2}),$$

and, as above, we obtain an estimate of the integral of this indicator function with respect to $P_{n,j}(dy)$ of order $C(\delta_{n-1} + n^{-1/2}||x||)$. Since we can use the third pseudomoment only, one needs to reduce the third derivative of g_i to the second one. Integrating by parts, we can rewrite the remainder term in (2.4) as

$$-\frac{1}{2}g_i''(y+n^{-1/2}x)[x^2]n^{-1} + \int_0^1 (1-\theta)g_i''(y+\theta n^{-1/2}x)[x^2]\,d\theta$$

and proceed as above. In this way we can prove bound (2.8) for $I_{j,2}$ and finally we arrive at the estimate

$$\Delta_{n,k,q}(r) \le C\varepsilon^{-3}n^{-1/2}\nu_3(\delta_{n-1}+\varepsilon) + C\varepsilon,$$

valid for all $r \ge 0$; thus, we have the following recursive inequality:

$$\delta_n \le C\varepsilon^{-3}n^{-1/2}\nu_3(\delta_{n-1}+\varepsilon).$$

In order to finish the proof we choose $\varepsilon = C_1\nu_3^{1/3}n^{-1/6}$ with an appropriate constant C_1 and use the induction assumption (2.2).

Theorem 2.1 is taken from Paulauskas and Račkauskas (1989). We refer the reader to this book for more information about the results obtained by Lindeberg's method combined with induction. Some generalizations of this approach are described there in detail.

The same method can be used to obtain nonuniform bounds on the convergence rate and to investigate probabilities of large deviations as well. This is possible due to the "good" estimates of the Gaussian measure of the set $\{r - \varepsilon \le q(x) \le r + \varepsilon\}$ for any continuous semi-norm q on B. For instance, when r.v. $q(Y)$ has a bounded density, a well-known result of Fernique yields for all $\lambda > 0$, $\varepsilon > 0$, $r \ge 0$

$$P\{r-\varepsilon \le q(Y) \le r+\varepsilon\} \le C\varepsilon\exp(-\lambda r),$$

where $C = C(\lambda, \mathcal{L}(Y))$. More precise results of this type are needed when considering large deviations. To this end the following lemma proved by Lifshits is quite useful (see Bentkus and Račkauskas (1990)).

Lemma 2.3. *Let q be a continuous semi-norm on B. Then*
(a) *there exists a constant $C = C(\mathcal{L}(Y)) > 0$ such that for all $\varepsilon > 0$, $r \geq 0$*

$$P\{q(Y) > r - \varepsilon\} \leq C \exp(Cr\varepsilon) P\{q(Y) > r\};$$

(b) *for any $r_0 > 0$ there exists a constant $C = C(r_0, \mathcal{L}(Y)) > 0$ such that for all $\varepsilon > 0$, $r \geq r_0$*

$$P\{r - \varepsilon \leq q(Y) \leq r + \varepsilon\} \leq C\varepsilon(r+1) P\{q(Y) > r - \varepsilon\}. \tag{2.9}$$

Moreover, (2.9) holds for all $r \geq 0$, $\varepsilon > 0$ with $C = C(\mathcal{L}(Y)) > 0$ if and only if the distribution of the r.e. $q(Y)$ has a bounded density.

Now, if we examine the formulas (2.6), (2.7) and make use of Lemma 2.3, we easily see the possibility of obtaining nonuniformity of power type in the bounds on $|I_{j,1}|$ and $|I_{j,2}|$. With some additional technical arguments the following result can be proved (for details, see Paulauskas and Račkauskas (1991)).

Theorem 2.4. *Let $m \geq 3$. Suppose that the conditions (A_3) and (D) are fulfilled. Then there exists a constant $C = C(m, \mathcal{L}(Y)) > 0$ such that*

$$\sup_{r \geq 0} r^m |P\{q(S_n(X)) < r\} - P\{q(Y) < r\}| \leq C \max(\nu_m n^{-m/2+1}, \nu_m^{1/m} n^{-1/6}).$$

Exponential nonuniformity is also available (see Paulauskas and Račkauskas (1991)). The validity of Lemma 2.3 in Hilbert space was pointed out by Bentkus (1986c), who investigated $P\{||S_n(X)|| > r\}$ for $0 \leq r \leq r_n$, where $r_n = O(n^{1/6})$. This result was generalized to Banach space and improved in Bentkus and Račkauskas (1990) by Lindeberg's method with induction and iteration arguments. A typical result is the following one.

Theorem 2.5 (Bentkus and Račkauskas (1990)). *Suppose that $q : B \to \mathbb{R}$ is continuous semi-norm and $E||X||^3 < \infty$. Suppose that the conditions (A_3), (D) are fulfilled. Then the following statements are equivalent:*

(a) *there exists $h > 0$ such that*

$$E \exp(hq^{1/2}(X)) < \infty;$$

(b) *there exist constants C_1, $C_2 < 0$ such that*

$$P\{q(S_n) > r\} \leq C_1 P\{q(Y) > r\}$$

for $0 \leq r \leq C_2 n^{1/6}$;
(c) *there exist constants C_1, $C_2 > 0$ such that*

$$P\{q(S_n) > r\} = P\{q(Y) > r\}(1 + \theta C_1 n^{-1/6}(1+r))$$

for $0 \leq r \leq C_2 n^{1/6}$, where $|\theta| \leq 1$;

(d) *for every function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) \to 0$ as $x \to \infty$,*

$$P\{q(S_n) > r\}/P\{q(Y) > r\} \to 1, \quad \text{as } n \to \infty$$

uniformly in r, $0 \leq r \leq f(n) n^{1/6}$.

Remarks on the proof. Assuming (a) to be true and employing Lemma 2.3 (instead of (2.7)) to help estimate $|I_{j,1}|$ and $|I_{j,2}|$, we obtain

$$|P\{q(S_n) > r\} - P\{q(Y) > r\}| \leq CP\{q(Y) > r\} n^{-1/8}(1+r)^{3/4}$$

when $0 \leq r \leq C_2 n^{1/6}$. This implies (b) which can be viewed as a first iteration for proving (c). The proof that (d) $\Rightarrow$ (b) proceeds by investigating the sequence $\sup[P\{q(S_n) > r\}/P\{q(Y) > r\} : r \leq f(n) n^{1/6}]$, $n \in \mathbb{N}$. One shows that if (b) is false, then this sequence fails to be bounded. To prove (b) $\Rightarrow$ (a), one applies Lévy's inequality.

As mentioned above, the bound given by Theorem 2.1 is in general sharp. The following result presents an example in Hilbert space.

Theorem 2.6 (Bentkus (1986b)). *Let a sequence $b_n \downarrow 0$. There exist a zero-mean Gaussian l_2-r.e. Y, a symmetric l_2-r.e. X and a continuous semi-norm q on l_2 such that*

(a) *$cov\, X = cov\, Y$, $P\{||X|| \leq 1\} = 1$;*
(b) *both conditions (A_3) and (D) are fulfilled;*
(c) *$\Delta_{n,q} \geq b_n n^{-1/6}$ for infinitely many n*

Sketch of the proof. Let I_s, $s \geq 1$, be blocks of natural numbers of length k_s, $s \geq 1$. Let $\lambda_i = 2^{-3i/2} k_i^{-2/3}$, $i \in \mathbb{N}$. Define

$$q(x) = \sum_{s=1}^{\infty} \lambda_s \sum_{k \in I_s} |x_k|.$$

The condition (A_3) for the semi-norm q can be deduced by means of the arguments of the proof of Theorem 2.2.31 in Paulauskas and Račkauskas (1989). To determine X and Y, we first define r.e.'s $X^{(j)}$, $j \in \mathbb{N}$, by setting

$$P\{X^{(j)} = a_j \lambda_j^{-1} e_i\} = P\{X^{(j)} = -a_j \lambda_j^{-1} e_i\} = (2k_j)^{-1}, \ i \in I_j,$$

where $\{e_i, i \in \mathbb{N}\}$ is the standard basis in l_2 and $a_j = 2^{-5j/3} k_j^{-2/3}$, $j \in \mathbb{N}$. Now let $X = \sum_{j=1}^{\infty} X^{(j)}$. It is easy to check that X is an l_2-r.e. The corresponding Gaussian l_2-r.e. is defined by $Y = \sum_{i=1}^{\infty} \tilde{\sigma}_i \gamma_i e_i$, where γ_i, $i \in \mathbb{N}$, is a sequence of i.i.d. standard normal r.v.'s and $\tilde{\sigma}_i = a_j \lambda_j^{-1} k_j^{-1/2}$ for $i \in I_j$, $j \in \mathbb{N}$. Since Y has independent coordinates, one can easily check that the distribution of r.v. $q(Y)$ has a bounded density. Therefore, the condition (D) follows. To prove (c) note, that

$$P\{q(S_n(X)) < r\} \geq P\{q(S_n(\tilde{X}^{(s)})) < r - \alpha_s\},$$

where $\tilde{X}^{(s)} = \sum_{i=1}^{s-1} X^{(i)}$, $\alpha_s = n^{1/2} \sum_{i=s}^{\infty} a_i$. Using the estimate of the rate of convergence in the CLT in $(k_1 + \cdots + k_{s-1})$-dimensional space (e.g., in Bentkus (1986a)), one deduces that

$$\Delta_{n,q} \geq \sup_{r\geq 0}[P\{q(\tilde{Y}^{(s)}) < r - \alpha_s\} - P\{q(Y) < r\}]$$
$$-C(k_1 + \cdots + k_{s-1})^2 n^{-1/2},$$

where $\tilde{Y}^{(s)} = \sum_{i=1}^{s-1} \sum_{k\in I_i} \tilde{\sigma}_k \gamma_k e_k$. To finish the proof one has to use the properties of the Gaussian r.e. Y and choose an appropriate sequence k_i, $i \in \mathbb{N}$.

In the case where one of the conditions (A_3), (D) fails to hold, the rate of convergence may be arbitrarily slow. Appropriate examples have been constructed by Bentkus (1984a), Rhee and Talagrand (1984) and Borisov (1985).

The verification of conditions (A_3), (D) is a separate problem. For details on the properties of the density of the r.v. $q(Y)$ for various functions $q : B \to \mathbb{R}$ we refer to the survey paper by Davydov and Lifshits (1985) and to Lifshits (1983) and Rhee and Talagrand (1986). Smooth approximation of an indicator function with various notions of smoothness are investigated in Bentkus and Račkauskas (1983) and Paulauskas and Račkauskas (1989), (1990).

We close this section with two results connected with conditions (A_3), (D). The modulus of convexity of a continuous semi-norm q on B is defined by

$$\tau_q(\varepsilon) = \inf\{1 - q(x - y/2) : q(x) = q(y) = 1,\ q(x - y) \geq \varepsilon\}.$$

Theorem 2.7 (Rhee and Talagrand (1986)). *If there exist constants $C > 0$, $\beta \geq 2$ such that $\tau_q(\varepsilon) \geq C\varepsilon^\beta$ for each $0 < \varepsilon \leq 2$, then the condition (D) holds.*

Note that the norm of the space L_p, $p > 1$, satisfies Theorem 2.7. The next result is usually used to construct a smooth approximation of an indicator function.

Theorem 2.8 (Bentkus and Račkauskas (1984)). *Condition (A_3) is equivalent to the following one. For each $\varepsilon > 0$ there exists a function $f_\varepsilon : B \to \mathbb{R}$ such that*

(a) *for all $x \in B$*

$$|q(x) - f_\varepsilon(x)| \leq \varepsilon;$$

(b) *the function f_ε is three times continuously Fréchet differentiable and*

$$\sup_{x\in B} ||f_\varepsilon^{(i)}(x)|| \leq C\varepsilon^{-i+1},\ i = 1, 2, 3.$$

§1.3. The Method of Integration by Parts

The method of integration by parts was introduced by Götze (1986). It allows one to obtain the Berry–Esseen-type bound for the speed of convergence in the CLT for sets $\{x \in B : F(x) < r\}$, $r \in \mathbb{R}$, where $F : B \to \mathbb{R}$ is a smooth function. When F is a linear function, this method reduces to Stein's method of differential equations. Further developments of Götze's approach are due to Zalesskii (1988). In order to describe the method we introduce the following conditions.

Differentiability condition (D_m). The function $F : B \to \mathbb{R}$ is m times Fréchet differentiable and there exist constants $C_F > 0$, $p \geq 0$ such that

$$||F^{(i)}(x)|| \leq C_F(1 + ||x||^p),\ i = 1, \ldots, m.$$

To formulate the next condition, let

$$\sigma^2(x) = E(F'(x)(Y))^2.$$

Variance condition (V). For any $c > 0$ and for sufficiently large $M > 0$ there exists a constant $C = C(c, M, \mathcal{L}(Y))$ such that for all $t \in \mathbb{R}$

$$\sup_{||a||\leq c} E\exp(-t^2\sigma^2(Y + a)) \leq C(1 + |t|)^{-M}.$$

Theorem 3.1. *If conditions (D_3), (V) are fulfilled and $E||X||^3 < \infty$, then*

$$\Delta_n := \sup_{r\in\mathbb{R}} |P\{F(S_n) < r\} - P\{F(Y) < r\}| = O(n^{-1/2}).$$

Götze (1986) proved Theorem 3.1 under the additional condition that the third derivative of F satisfies

$$||F'''(x + y) - F'''(y)|| \leq C_F(1 + ||x||^p + ||y||^p)||x||^\varepsilon$$

for some $\varepsilon > 0$, $p \geq 0$, and under a more complicated variance condition. Theorem 3.1 was proved by Zalesskii (1988). Actually Zalesskii proves a nonuniform result. To state it, let ρ denote the distance from $0 \in B$ to the boundary of the set $\{x \in B : F(x) < r\}$.

Theorem 3.2 (Zalesskii (1988)). *Suppose that conditions (D_3) and (V) hold and $E||X||^3 < \infty$. Then for any $s \geq 0$ there exists a constant $C = C(\mathcal{L}(X), s)$ such that*

$$|P\{F(S_n) < r\} - P\{F(Y) < r\}| \leq Cn^{-1/2}[(1 + \rho)^{-3} + (1 + |r|)^{-s}].$$

Now we give the main ideas of the proof of Theorem 3.1. For details, see the papers of Götze (1986) and Zalesskii (1988). The first step is to replace X_j by its truncation at the level $n^{1/2}$. One easily deduces that $\Delta_n \le \bar{\Delta}_n + Cn^{-1/2}$, where

$$\bar{\Delta}_n = \sup_{r\in\mathbb{R}} |P\{F(\bar{S}_n) < r\} - P\{F(Y) < r\}|,$$

$\bar{S}_n = n^{-1/2}(\bar{X}_1 + \cdots + \bar{X}_n)$; $\bar{X}_j = X_j$ if $\|X_j\| \le n^{1/2}$, and $\bar{X}_j = 0$ otherwise. Then one smoothes the indicator function of the set $\{F(x) < r\}$. To this end, let $f = f_{r,n} : \mathbb{R} \to \mathbb{R}$ be a monotone non-decreasing function such that $f(t) = 0$ if $t < r - n^{-1/2}$, $f(t) = 1$ if $t \ge r$, f is three times differentiable and $|f^{(i)}(t)| \le Cn^{i/2}$, $i = 1, 2, 3$. Define $f_1(x) = f(F(x))$, $f_2(x) = f(F(x) - n^{-1/2})$. Then we have

$$\bar{\Delta} \le \max_{i=1,2} \sup_{r\in\mathbb{R}} |I_i| + \sup_{r\in\mathbb{R}} P\{r \le F(Y) \le r + n^{-1/2}\},$$

where

$$I_i = I_i(r) = Ef_i(\bar{S}_n) - Ef_i(Y), \ i = 1, 2.$$

The next lemma proves the boundedness of the density of the r.v. $F(Y)$.

Lemma 3.3. *Let $m \ge 2$. If the function $F : B \to \mathbb{R}$ satisfies (D_{m+1}) and (V) and a function $\varphi : B \to \mathbb{R}$ satisfies (D_m), then for each $\varepsilon > 0$ and all $a \in \mathbb{R}$ there exists a constant $C = C(a, \varepsilon, \mathcal{L}(Y)) > 0$ such that for all $t \in \mathbb{R}$*

$$\sup_{\|x\|\le a} |E \exp(itF(Y + x))\varphi(Y + x)| \le C(1 + |t|)^{-m+\varepsilon}.$$

From Lemma 3.3 we easily obtain the estimate $\sup_{r\in\mathbb{R}} P\{r \le F(Y) \le r + \varepsilon\} \le C\varepsilon$. Therefore it remains to estimate I_i, $i = 1, 2$.

If identity (2.1) and a Taylor series are employed, then the third derivatives of the smoothed indicator function f would occur which are of order $O(n^{3/2})$. This is much too large to obtain an error bound of order $O(n^{-1/2})$. The main idea in integrating by parts is to insert factors in the integrals in I_i in such a way that integrating by parts helps to replace the derivatives $f^{(k)}$, $k = 1, 2, 3$, by the function f. To determine such factors, we need some preparation. Let X_{n+1}, X_{n+2}, X_{n+3} (respectively Y_{n+1}, Y_{n+2}, Y_{n+3}) denote independent copies of X (respectively Y). Let θ be a r.v. uniformly distributed on [0, 1]. Define random functions

$$g_{1,j}(x) = F'(x)[X_{n+j}], \ g_{2,j}(x) = F'(x)[Y_{n+j}], \ j = 1, 2, 3,$$

and functions

$$\sigma_1^2(x) = E(F'(x + \theta n^{-1/2}X)[X])^2, \ \sigma_2^2(x) = E(F'(x + \theta n^{-1/2}Y)[Y])^2.$$

Then $h_i(x) = \prod_{j=1}^3 g_{i,j}^2(x)(\sigma_i^2(x) + n^{-a})^{-3}$, $i = 1, 2$, $a > 1/2$, are the factors mentioned above. Put

$$\bar{U}_{k,n} = n^{-1/2}(\bar{X}_1 + \cdots + \bar{X}_k + Y_{k+1} + \cdots + Y_n);$$

and

$$Z_1 = t_1X_{n+1} + t_2X_{n+2} + t_3X_{n+3}, \ Z_2 = t_1Y_{n+1} + t_2Y_{n+2} + t_3Y_{n+3}.$$

Lemma 3.4. *For any $[n/2] \le k \le n$,*

$$|Ef(F(\bar{U}_{k,n})) - Ef(F(\bar{U}_{k,n} + Z_1))h_1(\bar{U}_{k,n} + Z_1)| = O(n^{-1/2}).$$

Lemma 3.5. *For any $1 \le k \le [n/2]$,*

$$|Ef(F(\bar{U}_{k,n})) - Ef(F(\bar{U}_{k,n} + Z_2))h_2(\bar{U}_{k,n} + Z_2)| = O(n^{-1/2}).$$

Now let us continue with the consequences of these results. We have

$$I_1 \le J_1 + J_2 + Cn^{-1/2},$$

where

$$\begin{aligned} J_1 &= |Eh(\bar{U}_{n,n_0} + Z_1) - Eh(\bar{U}_{n,n} + Z_1)|, \\ J_2 &= |Eh(\bar{U}_{n,n_0} + Z_2) - Eh(\bar{U}_{n,0} + Z_2)|, \end{aligned}$$

$h(z) = f(F(z))h_1(z)$ and n_0 is the integer $[n/2]$.

Now one proceeds as usual. One uses the identity (2.1) and then expands the function h in a Taylor series up to terms of second order obtaining, e.g.,

$$J_1 \le \sum_{k=n_0}^{n} [|J_{1,k}^{(1)}| + |J_{1,k}^{(2)}| + |J_{1,k}^{(3)}|],$$

where

$$\begin{aligned} J_{1,k}^{(1)} &= \int_B \int_B \chi(\|x\| \ge \sqrt{n})h(y)Q_{k,n}(dy)H(dx); \\ J_{1,k}^{(2)} &= \int_B \int_B \chi(\|x\| \ge \sqrt{n})h'(y)[x]n^{-1/2}Q_{k,n}(dy)H(dx); \\ J_{1,k}^{(3)} &= \int_B \int_B h''(y + \lambda n^{-1/2}x)[x^2]n^{-1}Q_{k,n}(dy)\bar{H}(dx), \ \lambda \in [0, 1], \end{aligned}$$

$$H = \mathcal{L}(X) - \mathcal{L}(Y), \ \bar{H} = \mathcal{L}(\bar{X}) - \mathcal{L}(Y),$$

$$Q_{n,k} = \mathcal{L}(\bar{W}_{k,n}), \ \bar{W}_{n,k} = n^{-1/2}\left(\sum_{i=1}^{k-1} \bar{X}_i + \sum_{i=k+1}^{n} Y_i\right).$$

The following lemma is important for the further estimations. Put

$$\sigma_0^2(x) = E(F'(x)[\bar{X}_1])^2.$$

Lemma 3.6. *For any $c > 1/2$, $b > 0$, $a \in \mathbb{R}$ there exists a positive constant C such that for any $k = 1, 2, \ldots, n$*

$$\sup_{\|x\| \le a} E(\sigma_0^2(\bar{W}_{n,k} + x) + n^{-c})^{-b} \le C.$$

Integration by parts is applied after rewriting

$$g_{1,1}(y + Z_1) f'(F(y + Z_1)) = \frac{\partial}{\partial t_1} f(F(y + Z_1));$$

$$g_{1,1}(y + Z_1) g_{1,2}(y + Z_1) f''(F(y + Z_1)) = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2} f(F(y + Z_1)),$$

etc., and using the fact that the factors $g_{1,j}(y+Z_1)$ are always attached to the derivatives of the function f. Now one considers $J^{(i)}_{1,k} = \int_0^1 \int_0^1 \int_0^1 J^{(i)}_{1,k}\, dt_1\, dt_2\, dt_3$ and integrates by parts to estimate these integrals.

Integration by parts is also useful when investigating the probabilities of large deviations. As an illustration, we state the following result (cf. Theorem 2.5 and Theorem 1.7).

Theorem 3.7 (Račkauskas (1988)). *Let X, Y be H-valued r.e.'s. The following statements are equivalent:*

(a) *there exists $\lambda > 0$ such that*

$$E \exp(\lambda \|X\|^{1/2}) < \infty;$$

(b) *there exist constants $C_i = C(\mathcal{L}(Y)) > 0$, $i = 1, 2$, such that*

$$P\{\|S_n\| > r\} = P\{\|Y\| > r\}(1 + C_1 \theta (1 + r) n^{-1/2}),$$

when $0 \le r \le C_2 n^{1/6}$. Here $|\theta| \le 1$.

Earlier Zalesskii (1989) proved (b) under the slightly stronger moment condition $E \exp(\lambda \|X\|) < \infty$.

§1.4. The Method of Finite-Dimensional Approximation

This section discusses the method of finite-dimensional approximation in the context of estimating the convergence rates in limit theorems in infinite-dimensional spaces. At the end of the seventies, it seemed that this method yields less accurate results compared with other methods. The first estimate in the CLT in infinite-dimensional Hilbert space, due to Kandelaki (1965) (see also Vakhaniya and Kandelaki (1969)), by means of finite-dimensional approximation, had only an inverse logarithmic order of decay. Sazonov (1968, 1969) used finite-dimensional approximation in the CLT in Hilbert space for summands having a special structure (for the ω^2-test) and he achieved the rate

$O(n^{-1/6+\varepsilon})$, $\varepsilon > 0$. Giné (1976) and Paulauskas (1976a) used this method to find the convergence rate in the CLT in $C(S)$, where S is a compact metric space. For $S = [0, 1]$ and the limiting Wiener process, they obtained the rate $n^{-1/20}$ or even less. These papers studied the special compact set $S = \{0, 1/2, 1/3, \ldots\}$ under very strict conditions on the terms: sub-Gaussian increments and a very slow rate of growth of the metric entropy, etc. The rate found in these papers was $n^{-1/6+\varepsilon}$, $\varepsilon > 0$. In a later paper, Paulauskas (1984) came up with $n^{-1/2+\varepsilon}$, $\varepsilon > 0$.

There are more papers where the method of finite-dimensional approximation has been applied under different settings. But a general feature of all is the rather slow rate of the bounds obtained. This can be explained by the fact that the estimates such as

$$\sup_{A \in \mathcal{A}} |P\{S_n(X) \in A\} - P\{Y \in A\}| \le C(k) C(\mathcal{L}(X), \mathcal{L}(Y)) n^{-1/2}$$

on some class $\mathcal{A}$ of Borel sets in $\mathbb{R}^k$ (for example, the class of rectangles or the class of convex sets) depend rather heavily on the dimension k. The present estimates of the constant $C(k)$ are of the form $C(k) \le Ck^\beta$ for some absolute constant C and some exponent β. They depend on the form of the constant $C(\mathcal{L}(X), \mathcal{L}(Y))$ and on the class of sets $\mathcal{A}$ under consideration. For the class of convex sets and the identity covariance operator of X, Nagaev (1976) and Senatov (1983) proved the estimate with $\beta = 1$ and $C(\mathcal{L}(X), \mathcal{L}(Y)) = E\|X\|^3$. Bentkus (1986a) improved this result to $\beta = 1/2$. We note that for the method of finite-dimensional approximation usually one needs estimates in $\mathbb{R}^k$ over the class of rectangles only without assumptions about the covariance structure of random vectors. To this end, there is a useful result due to Bentkus (1984c, 1990). Namely, he constructed smooth functions approximating well the indicator functions of balls in l^k_∞ with precise estimates of the derivatives of these functions (see Lemma 4.6 below). This allowed him to obtain new bounds on the remainder term in the CLT in l^k_∞ for the class of balls with worse dependence on n but with much better dependence on k, namely, logarithmic. This resulted in a better convergence rate in the CLT in the space $C(S)$, S a compact metric space (see Bentkus (1982) and Paulauskas and Račkauskas (1989) for full information on this topic) and even in other infinite-dimensional spaces (see Paulauskas and Juknevičienė (1988) for a generalization to the Skorokhod space $D[0, 1]$ and Norvaiša and Paulauskas (1990) for the case of general empirical processes).

Now we describe this method in detail. Suppose that our r.e.'s X, X_1, X_2, … are random processes, defined on a probability space $(\Omega, \mathcal{A}, P)$ and indexed by some parameter set T, which we assume to be a compact metric set, i.e., $X : \Omega \times T \to \mathbb{R}$. We shall view X as a map from Ω to $l_\infty(T)$, the space of all bounded functions on T with supremum norm or to some smaller subspace such as $C(T)$. For $x \in l_\infty(T)$, $\|x\|_\infty \equiv \|x\|_{T,\infty} := \sup_{t \in T} |x(t)|$; let $EX(t) = 0$ and $EX^2(t) < \infty$ for all $t \in T$. Suppose that the corresponding Gaussian random element $Y = \{Y(t), t \in T\}$ with $EY(t) = 0$ and

$EY(t)Y(s) = EX(t)X(s)$ for all $s, t \in T$, is in $l_\infty(T)$ (or in a smaller subspace if X is concentrated in it). Finally, (in order not to use outer measure and integrals) suppose that both $||S_n(X)||_\infty$ and $||Y||_\infty$ are measurable. Therefore the quantities

$$\Delta_n(r) = \Delta_n(T,r) = |P\{||S_n(X)||_\infty < r\} - P\{||Y||_\infty < r\}|, \tag{4.1}$$

$$\Delta_n = \sup_{r\geq 0} \Delta_n(T,r)$$

are well defined. Before proceeding to estimate (4.1), we give the main examples which fit into this framework.

(a) $T = [a,b]$, X and Y, as processes on $[a,b]$, are a.s. continuous. Here an estimate of $\Delta_n(r)$ yields the rate of convergence in the CLT in $C[a,b]$.

(b) $T = [a,b]$, X and Y have no discontinuities of the second kind a.s. Here $\Delta_n(r)$ measures the rate of convergence in the CLT in $D[a,b]$ on balls with respect to the supremum norm. Later on we shall consider this example in detail (see §3.3).

(c) It is possible to generalize (b) and to consider the space $D(\mathbb{A})$ of functions $x : \mathbb{A} \to \mathbb{R}$, indexed by some class $\mathbb{A}$ of closed Borel subsets of $I^d (\equiv [0,1]^d)$, which are outer continuous and have inner limits. This space was introduced by Bass and Pyke (1985, 1987) where general limit theorems for a triangular array were proved. If one takes $T = \mathbb{A}$ in (4.1), it should be interesting to find estimates of the convergence rate in these general theorems.

(d) Let $Z, Z_1, Z_2, \ldots$ be i.i.d. r.v.'s with values in a measurable space $(\chi, \mathcal{A})$ and with a distribution m. The empirical process associated with m is given by

$$E_n = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{Z_i} - m\right).$$

We consider the process $E_n(f)$, $f \in \mathcal{F}$, indexed by some class $\mathcal{F} \subset L_2(\chi, \mathcal{A}, m)$ of real-valued measurable functions $f : \chi \to \mathbb{R}$. Here and in what follows we put $\mu(f) = \int_\chi f(x)\mu(dx)$. In particular, the class $\mathcal{F}$ can consist of the indicator functions of some class of sets. Let $\{G_m(f), f \in \mathcal{F}\}$ denote the limiting Gaussian process (see Giné and Zinn (1984) for details). For simplicity, assume that $\mathcal{F}$ is countably generated (for m), that is, there exists a countable subclass $\mathcal{F}_0 \subset \mathcal{F}$ such that $\sup_{f\in\mathcal{F}} |E_n(f)| = \sup_{f\in\mathcal{F}_0} |E_n(f)|$ a.s. for all $n \geq 1$. A possible way to estimate the convergence rate in the CLT for empirical processes $E_n(f)$, $f \in \mathcal{F}$, is to bound the quantity

$$\delta_n(\mathcal{F}, r) = |P\{\sup_{f\in\mathcal{F}} |E_n(f)| < r\} - P\{\sup_{f\in\mathcal{F}} |G_m(f)| < r\}|.$$

This coincides with $\Delta_n(T,r)$ in (4.1) if we choose $X = \delta_Z - m$ and $T = \mathcal{F}$. Rates of convergence for empirical processes will be described in the last chapter. This important example has an additional theoretical aspect. It is known that in any Banach space an estimate in the CLT on balls can be obtained by estimating the remainder term in the CLT for the empirical process with class $\mathcal{F}$ being the unit sphere of the dual space B^* (due to the fact $||x|| = \sup\{|f(x)| : f \in B^*,\ ||f|| = 1\}$).

We now return to estimating the quantity $\Delta_n(T,r)$ in (4.1). Suppose that ϱ is some pseudometric on T under which T is totally bounded. Usually ϱ is connected with the process X. The pseudometric $\tau(s,t) = E^{1/2}(X(s) - X(t))^2$, $s, t \in T$, provides such an example. Let $\delta > 0$ be arbitrary and $N = N_\varrho(\delta)$ be the number of elements of a minimal δ-net $T(\delta) = \{t_1, t_2, \ldots, t_N\}$ of the totally bounded set (T, ϱ). For $x \in l_\infty(T)$ and $\delta > 0$, put

$$w_\varrho(x,\delta) = \sup\{|x(t) - x(s)| : t, s \in T,\ \varrho(t,s) < \delta\};$$

$$||x||_N := \max_{t\in T(\delta)} |x(t)| = \max_{1\leq i\leq N} |x(t_i)|;$$

$$\Delta_n^N(r) = \Delta_n(T(\delta), r) = |P\{||S_n||_N > r\} - P\{||Y||_N > r\}|.$$

Obviously, for each n and any $0 < \varepsilon < r$, we have

$$\begin{aligned}
&P\{||S_n||_\infty > r\} - P\{||Y||_\infty > r\} = P\{||S_n||_\infty > r\} - P\{||S_n||_N > r - \varepsilon\} \\
&\quad + P\{||S_n||_N > r - \varepsilon\} - P\{||Y||_N > r - \varepsilon\} \\
&\quad + P\{||Y||_N > r - \varepsilon\} - P\{||Y||_\infty > r\} \\
&\leq P\{||S_n||_\infty > r, ||S_n||_N \leq r - \varepsilon\} \\
&\quad + \Delta_n^N(r - \varepsilon) + P\{||Y||_N > r - \varepsilon, ||Y||_\infty \leq r\} \\
&\leq P\{w_\varrho(S_n, \delta) > \varepsilon, ||S_n||_\infty > r\} \\
&\quad + \Delta_n^N(r - \varepsilon) + P\{r - \varepsilon \leq ||Y||_\infty \leq r\}.
\end{aligned}$$

Analogously we have the lower bound

$$\begin{aligned}
&P\{||S_n||_\infty > r\} - P\{||Y||_\infty > r\} \\
&\geq -\Delta_n^N(r) - P\{r \leq ||Y||_\infty \leq r + \varepsilon\} - P\{w_\varrho(Y,\delta) > \varepsilon, ||Y||_\infty > r\}.
\end{aligned}$$

These estimates lead to the following lemma which presents the core of the method of finite-dimensional approximation.

Lemma 4.1. *For all n and all $\varepsilon > 0$, $r > \varepsilon$, $\delta > 0$,*

$$\begin{aligned}
&\Delta_n(T,r) \leq \Delta_n^N(r - \varepsilon) + \Delta_n^N(r) \\
&+ P\{w_\varrho(S_n,\delta) > \varepsilon, ||S_n||_\infty > r\} + P\{w_\varrho(Y,\delta) > \varepsilon, ||Y||_\infty > r\} \\
&+ P\{r - \varepsilon \leq ||Y||_\infty \leq r + \varepsilon\}.
\end{aligned} \tag{4.2}$$

Remark 4.2. The last term in (4.2) is the Gaussian measure of an ε-strip straddling the boundary of a ball in $l_\infty(T)$. Such terms appear in all methods used to estimate the convergence rate (see §§1.1, 1.2). In what follows we assume additionally that for each $m > 0$ there exists a constant $C = C(\mathcal{L}(Y), m)$ such that for all $\varepsilon > 0$, $r > 0$

$$P\{r \leq ||Y||_\infty \leq r + \varepsilon\} \leq C\varepsilon(1 + r)^{-m}. \tag{4.3}$$

Remark 4.3. The bound (4.2) has been proved for $r > \varepsilon$. If we want to obtain a general estimate of $\Delta_n(T,r)$ involving uniform and nonuniform bounds simultaneously we need to estimate $\sup_{r\geq 0}(1+r)^m\Delta_n(T,r)$. It is easy to see that this can be reduced to estimating over $r > \varepsilon$ only. Indeed, utilizing (4.3), for any $0 < r \leq \varepsilon$, we have

$$\begin{aligned}(1+r)^m\Delta_n(T,r) &\leq (1+\varepsilon)^m\Delta_n(T,\varepsilon) + 2P\{\|Y\|_\infty \leq \varepsilon\}\\ &\leq \sup_{r>\varepsilon}(1+r)^m\Delta_n(T,r) + C\varepsilon.\end{aligned}$$

Remark 4.4. Sometimes it is better to bound $P\{\|S_n\|_\infty > r, \|S_n\|_N \leq r-\varepsilon\}$ by

$$P\{\sup_{1\leq i\leq N}\sup_{\varrho(t,t_i)<\delta}|S_n(t) - S_n(t_i)| > \varepsilon, \|S_n\|_\infty > r\}.$$

We now discuss briefly the estimation of the terms in (4.2), except for the last one, for which we require (4.3). The terms

$$P\{w_\varrho(S_n,\delta) > \varepsilon, \|S_n\|_\infty > r\},\ P\{w_\varrho(Y,\delta) > \varepsilon, \|Y\|_\infty > r\}$$

in (4.2) control the oscillation of the processes S_n and Y, respectively. Using the elementary inequality

$$P\{A \cap B\} \leq (P\{A\})^\gamma(P\{B\})^{1-\gamma},$$

valid for all $0 \leq \gamma \leq 1$, we reduce the estimation of these quantities to that of

$$P\{w_\varrho(S_n,\delta) > \varepsilon\},\ P\{w_\varrho(Y,\delta) > \varepsilon\}, \tag{4.4}$$

$$P\{\|S_n\|_\infty > r\},\ P\{\|Y\|_\infty > r\}. \tag{4.5}$$

To bound the terms in (4.5) the Chebyshev inequality suffices. Since we are assuming that the Gaussian process Y, as a process on the parameter set T, is bounded, we can use an exponential bound for the tail of the distribution of the Gaussian process (see Fernique (1971) and Marcus and Shepp (1972)): there exists a finite positive constant C such that for all $\varepsilon, \delta \in \mathbb{R}^+$

$$P\{w_\varrho(Y,\delta) > \varepsilon\} \leq C\exp(-\varepsilon^2/4\sigma_\varrho^2(\delta)),$$

where $\sigma_\varrho^2(\delta) = \sup\{E(Y(t)-Y(s))^2 : \varrho(s,t) \leq \delta\}$. The only term for which we are unable to recommend a general approach is $P\{w_\varrho(S_n,\delta) > \varepsilon\}$ since under different assumptions on $\mathcal{L}(X)$ different bounds for this term are available.

The estimation of $\Delta_n^N(r)$ and $\Delta_n^N(r-\varepsilon)$ is equivalent to estimating the convergence speed in the CLT in the finite-dimensional Banach space l_∞^N. In estimating $\Delta_n^N(r)$, one must keep in mind that N (through its dependence on $\delta = \delta_n \downarrow 0$) will grow with n. Another feature in estimating $\Delta_n^N(r)$ is the appearance of the probability

$$P\{r \leq \max_{1\leq i\leq N}|Y(t_i)| \leq r+\varepsilon\}. \tag{4.6}$$

We can bound the quantity (4.6) by $C\varepsilon$ using (4.3) and the term controlling the oscillation of the process Y.

We shall state an estimate for $\Delta_n^N(r)$ when $E\|X\|_\infty^3 < \infty$. Write

$$\begin{aligned}W_\varrho(\mathcal{L}(Y),\delta,t) &:= \sup_{r\geq 0}(1+r)^3P\{w_\varrho(Y,\delta) > t, \|Y\|_\infty > r\},\\ M_3(n) &= n^{-1/6}(1 \vee (E\|X\|_\infty^3 + E\|Y\|_\infty^3)^{1/3}),\\ D_n(\delta) &= M_3(n) \vee W_\varrho(\mathcal{L}(Y),\delta,M_3(n)),\\ H_\varrho(u) &= \log N_\varrho(u),\\ a\vee b &= \max(a,b),\ a\wedge b = \min(a,b).\end{aligned}$$

Theorem 4.5. *There exists an absolute constant $C < \infty$ such that for all $\delta > 0$, $n \in \mathbb{N}$, $r \geq 0$ the following estimate holds:*

$$\Delta_n^{N_n}(r) \leq C(1+r)^{-3}H_\varrho^2(\delta)(D_n(\delta) \vee D_n^3(\delta)), \tag{4.7}$$

where $N_n = N_\varrho(\delta)$.

This result is taken from Norvaiša and Paulauskas (1990). Earlier results of this kind (roughly speaking, giving an order $n^{-1/6}$ for Δ_n) were proved by Paulauskas and Račkauskas (1989) (see Theorem 5.2.6), and Paulauskas and Juknevičienė (1988). In the formulation of (4.7) we did not seek the best possible accuracy. It is easy to see that instead of the moments $E\|X\|_\infty^3 + E\|Y\|_\infty^3$ in (4.7) one can use

$$\int \sup_{t\in T}|x(t)|^3|\mathcal{L}(X) - \mathcal{L}(Y)|(dx).$$

But in estimating the oscillation of the processes S_n and Y, at present one cannot avoid using the moments $E\|X\|_\infty^p$ and $E\|Y\|_\infty^p$, $p > 3$. Therefore we had to restrict ourselves to (4.7). A more important question concerns the power with which the metric entropy $H_\varrho(\delta_n)$ occurs in (4.7). It is possible to show that this power can be lowered from 2 to 2/3 in two cases: if we confine ourselves to uniform estimates of $\Delta_n^N = \sup_{r\geq 0}\Delta_n^N(r)$ or if we allow $\sup_n E\|S_n\|_\infty^2$ to appear in the final estimate of $\Delta_n(T,r)$.

The proof of Theorem 4.5 runs along the lines of Lindeberg's method, described in the previous section. Therefore we shall give only a sketch of the proof. Define

$$\xi_j = (X_j(t_1),\ldots,X_j(t_N)),\ \eta_j = (Y_j(t_1),\ldots,Y_j(t_N)),$$

where Y, Y_1, Y_2, ... are i.i.d. r.e.'s. Instead of $\Delta_n^N(r)$ it is more convenient to estimate the quantity

$$\begin{aligned}\delta_n^N = \sup_{1\leq j\leq n}\sup_{r\geq 0} r^3|P\{n^{-1/}\|\xi_1 + \cdots + \xi_j + \eta_{j+1} + \cdots + \eta_n\|_N \leq r\}\\ -P\{\|\eta\|_N \leq r\}|,\end{aligned}$$

which majorizes $\Delta_n^N(r)r^3$. The first step is usually called smoothing and uses the following result due to Bentkus (1990) (see also Bentkus (1984c)).

Lemma 4.6. *For all $r \geq 0$, $\varepsilon > 0$ there exists a function $f_{r,\varepsilon} : l_\infty^N \to [0,1]$, $f_{r,\varepsilon} \in C^\infty$, such that*

$$\chi\{||x||_N \leq r\} \leq f_{r,\varepsilon}(x) \leq \chi\{||x||_N \leq r+\varepsilon\},$$

$$||f_{r,\varepsilon}^{(m)}(x)|| \leq C(m)\varepsilon^{-m}\ln^{m-1}(N+1), \; m = 1, 2, \ldots.$$

Furthermore, the constant $C(m)$ depends on m only.

The further steps in the proof are standard: one uses the identity (2.1) and the Taylor expansion up to the third order. The only change is that one has to apply the estimate

$$P\{r-\varepsilon \leq ||\eta||_N \leq r+\varepsilon\} \leq$$
$$P\{r-\varepsilon-t \leq ||Y||_\infty \leq r+\varepsilon+t\} + P\{w_\varrho(Y,\delta) > t, ||Y||_\infty > r-\varepsilon\}$$

for any $t \geq 0$. This results in the following recursive inequality:

$$\delta_n^N \leq M_3^3(n)\left\{C + CH_\varrho^2(\delta)t^{-3}\left[\delta_{n-1}^{N_n} + t + W_\varrho(Y,\delta,t)\right]\right\}$$
$$+C[t + t^3 + W_\varrho(Y,\delta,t)].$$

A standard induction argument completes the estimation of $\delta_n^N \geq r^3\Delta_n^N(r)$. The bound for $\sup_{r\geq 0}\Delta_n^N(r)$ is obtained in an analogous way.

It is worth mentioning that the method of finite-dimensional approximation can be used not only for the supremum norm. For example, if we want to estimate the remainder term in the CLT in the space l_p, $1 \leq p < \infty$, on balls, then as a finite-dimensional approximation it is natural to take the first N coordinates and to estimate the remainder term in the CLT in l_p^N for balls, combining this with estimates of the tails of the coordinates. Such an approach has been used by Paulauskas (1981) and in l_2 by Sazonov (1968, 1969).

In the paper due to Asriev and Rotar' (1985) an estimate in l_∞^k for parallelepipeds was obtained (having the order $n^{-1/2}$ with a logarithmic factor). The estimate is expressed in such a form that it allows one to pass to the limit as $k \to \infty$ and to get a bound (of the same order) in $\mathbb{R}^N$. Unfortunately we are unable to apply this result in the context of finite-dimensional approximation since it is obtained under the assumption that the covariance matrix of the random vector X under consideration is diagonal. This assumption is not restrictive at all in $\mathbb{R}^k$ in the case of the class of convex sets since by means of an orthogonal transformation, the covariance matrix can be diagonalized. But this is impossible for the class of parallelepipeds with sides parallel to the coordinate axes.

We shall now demonstrate how the general scheme of finite-dimensional approximation applies to the CLT in the Skorokhod space $D[0,1]$. But before we do this we shall give a short (and therefore not full) review of what is known about the CLT in the space $D[0,1]$ itself, since this topic is not covered in the literature so thoroughly as the CLT in Banach spaces (see, for example, Araujo and Giné (1980) and Paulauskas and Račkauskas (1989)).

Let $X, X_1, X_2, \ldots$ be i.i.d. r.e.'s with values in $D[0,1]$. We assume that $D[0,1]$ is equipped with the Skorokhod topology and metric under which it is a separable and complete metric space. Let us assume that $EX(t) = 0$, $EX^2(t) < \infty$, for all $t \in [0,1]$. We say that X satisfies the CLT in $D[0,1]$ ($X \in CLT(D)$ for short) if there exists a Gaussian zero-mean $D[0,1]$-valued r.e. Y such that $EX(t)X(s) = EY(t)Y(s)$ for all $s,t \in [0,1]$ and $S_n = n^{-1/2}\sum_{i=1}^n X_i$ converges in distribution to Y (see, e.g., Billingsley (1968)). The CLT in D was considered by Fisz (1959), Hahn (1978), Bass and Pyke (1985), Juknevičienė (1985), Paulauskas and Stieve (1990), and Bézandry and Fernique (1990). Many applied problems lead directly to the CLT in $D[0,1]$. One such example is in a paper of Phoenix and Taylor (1973) where the asymptotic strength distribution of a general fibre bundle was investigated. As a matter of fact this investigation goes back to an early paper of Daniels (1945). Pheonix and Taylor (1973) did prove the CLT in $D[0,1]$ (without stating it explicitly for i.i.d. random processes with a special structure. Influenced by this result, Hahn (1978) proved the following theorem.

Theorem 4.7. *Let X be a r.e. in $D[0,1]$, $EX(t) = 0$ and $EX^2(t) < \infty$ for all $t \in [0,1]$. Assume that there exist nondecreasing continuous functions F_1 and F_2 on $[0,1]$ and numbers $\alpha_1 > 1/2$, $\alpha_2 > 1$ such that for all $0 \leq s \leq t \leq u \leq 1$ the following two inequalities hold:*

$$E(X(t)-X(s))^2 \leq (F_1(t)-F_1(s))^{\alpha_1}, \tag{4.8}$$

$$E(X(u)-X(t))^2(X(t)-X(s))^2 \leq (F_2(u)-F_2(s))^{\alpha_2}. \tag{4.9}$$

Then $X \in CLT(D)$ and $\mathcal{L}(Y)(C[0,1]) = 1$, where Y is the limiting Gaussian process for $S_n(X)$.

Condition (4.9) requires the finiteness of the fourth moment $EX^4(s)$ for all $s \in [0,1]$. This shortcoming is eliminated in the following theorems.

Theorem 4.8 (Paulauskas and Stieve (1990)). *Suppose that X is a r.e. in $D[0,1]$ and $EX(t) = 0$, $EX^2(t) < \infty$ for all $t \in [0,1]$. Suppose that (4.8) is satisfied for some $\alpha_1 > 2/3$ and the following inequality holds for some $\alpha_2 > 1$:*

$$E\left((X(t)-X(s))^2 \wedge 1\right)\left(X(u)-X(t)\right)^2 \leq \left(F_2(u)-F_2(s)\right)^{\alpha_2}. \tag{4.10}$$

Then $X \in CLT(D)$ and $\mathcal{L}(Y)(C[0,1]) = 1$.

Theorem 4.9 (Bézandry and Fernique (1990)). *Let X be a real-valued random function on $[0,1]$, defined on some probability space $(\Omega, \mathcal{A}, P)$. Suppose that there exist continuous increasing functions δ, η, and θ from $[0,1]$ to $\mathbb{R}$ such that θ is concave and $\delta(0) = \eta(0) = \theta(0) = 0$. Moreover, let the following conditions hold for all $0 \leq s \leq t \leq u \leq 1$ and all $A \in \mathcal{A}$:*

$$E|X(0)|^2 < \infty, \; E|X(t)-X(s)|^2 \leq \delta^2(t-s), \tag{4.11}$$

$$\int_0^1 u^{-5/4}(\log(1+1/u))^{1/4}\delta(u)\,du < \infty, \tag{4.12}$$

$$E(|X(s)-X(t)|^2 \wedge |X(t)-X(u)|^2)\chi_A \le \eta^2(u-s)\theta(P(A)), \tag{4.13}$$

$$\int_0^1 u^{-3/2}\theta^{1/2}(u\log_2(1+1/u))\eta(u)\,du < \infty. \tag{4.14}$$

Then $X \in CLT(D)$ *and* $\mathcal{L}(Y)(C[0,1]) = 1$.

It seems that Theorems 4.7–4.9 are not comparable. It is worth mentioning that if we require the boundedness of a random process X, then a very mild condition on the increments of the process is sufficient for the CLT. This is demonstrated in the following result.

Theorem 4.10 (Giné and Zinn (1984)). *Let* X *be a centered stochastically continuous and uniformly bounded process with sample functions in* $D[0,1]$. *Assume that there exist a positive* C *and a nondecreasing function* $F \in D[0,1]$ *such that for all* $s,t \in [0,1]$

$$E|X(t)-X(s)| \le C|F(t)-F(s)|.$$

Then $X \in CLT(D)$. *Furthermore, if* $\mathcal{L}(X)(C[0,1]) = 1$, *then*

$$X \in CLT(C[0,1]).$$

In fact, this theorem is a consequence of a more general result which, in turn, is derived from the general CLT for empirical processes (see Giné and Zinn (1984)).

We mention that another result (also apparently uncomparable with others) on the CLT in $D[0,1]$ can be derived from Bass and Pyke (1987).

Remark 4.11. A final result on the CLT in $D[0,1]$ for i.i.d. summands, formulated in terms of moments of increments of a process, was obtained independently by Bloznelis and Paulauskas (1994) and Fernique (1994).

Next we shall state and give a sketch of the proof of a result due to Paulauskas and Stieve (1990), which provides an estimate of the convergence rate in Theorem 4.8. As usual, the modulus of continuity of a function f is defined by $\omega_f(\delta) = \sup\{|f(t)-f(s)| : |s-t| < \delta\}$.

Theorem 4.12. *Let* X *satisfy the conditions of Theorem 4.8. Additionally, assume that there exist a nondecreasing continuous function* F_3, *numbers* $\alpha_3 \ge 1$, β_1, β_2, $\beta_3 > 0$ *and* $0 \le \kappa \le 1$ *such that for all* $0 \le s \le t \le 1$

$$E|X(t)-X(s)|^{3+\kappa} \le (F_3(t)-F_3(s))^{\alpha_3}, \tag{4.15}$$

$$\omega_{F_i}(\delta) \le C\delta^{\beta_i},\ i = 1,2,3. \tag{4.16}$$

Suppose that $E\sup_{t\in[0,1]}|X(t)|^3 < \infty$ *and condition (4.3) holds with* $T = [0,1]$. *Then there exists a finite constant* $C = C(m, \mathcal{L}(X), \mathcal{L}(Y))$ *such that*

$$\begin{aligned}\Delta_n &:= |P\{||S_n(X)||_\infty < \lambda\} - P\{||Y||_\infty < \lambda\}| \\ &\le Cn^{-\varphi(\kappa)}(1+\lambda)^{-3}(\ln n + \ln(1+\lambda))^2,\end{aligned} \tag{4.17}$$

where

$$\varphi(\kappa) = \begin{cases} 1/6 & \text{if } \alpha_3 > 1; \\ 1/6 \vee (1+\kappa)/10 & \text{if } \alpha_3 = 1. \end{cases}$$

Sketch of the proof. To avoid certain technicalities, we shall sketch the proof under stronger conditions then those stated in the theorem. Namely, we shall assume instead of (4.10) the conditions (4.9) and (4.15) with $\kappa = 1$ (therefore in (4.8) we can assume that $\alpha_1 > 1/2$). Under such conditions an estimate of the remainder term was given in Paulauskas and Juknevičienė (1988). Moreover, we assume that $\alpha_3 > 1$ and $E\sup_{t\in[0,1]}|X(t)|^p < \infty$ for some $p > 3$. Let $0 < \delta < 1$ be a number to be specified later on and let $N = [1/\delta]+1$ and $0 = t_1 < t_2 < \cdots < t_N = 1$ be such that $t_k - t_{k-1} \le \delta$. As earlier $||x||_N = \sup_{1\le i\le N}|x(t_i)|$ and subscript N refers to the corresponding quantities in the space l_∞^N. Applying Lemma 4.1 and Remark 4.4, for any $0 \le \varepsilon \le \lambda$ and $0 \le \delta < 1$, we have the estimate

$$\Delta_n(\lambda) \le I_1 + I_2 + I_3 + \Delta_n^N(\lambda-\varepsilon) + \Delta_n^N(\lambda), \tag{4.18}$$

where

$$\begin{aligned} I_1 &= P\{\sup_{1\le i\le N-1}\sup_{t_i\le t\le t_{i+1}}|S_n(t)-S_n(t_i)| > \varepsilon,\ ||S_n||_\infty > \lambda\}, \\ I_2 &= P\{\sup_{1\le i\le N-1}\sup_{t_i\le t\le t_{i+1}}|Y(t)-Y(t_i)| > \varepsilon, ||Y||_\infty > \lambda\}, \\ I_3 &= P\{\lambda-\varepsilon \le ||Y||_\infty \le \lambda+\varepsilon\}. \end{aligned}$$

By some straightforward computations, we arrive at having to estimate the quantities

$$\begin{aligned} I_{1,1} &= P\{\sup_{1\le i\le N-1}\sup_{t_i\le t\le t_{i+1}}|S_n(t)-S_n(t_i)| > \varepsilon\}, \\ I_{2,1} &= P\{\sup_{|t-s|<\delta}|Y(t)-Y(s)| > \varepsilon\}. \end{aligned}$$

To estimate $I_{2,1}$ it is sufficient to use the following result (see, for example, Marcus and Pisier (1981)).

Lemma 4.13. *Let* η *be a zero-mean Gaussian* $C[0,1]$*-valued r.e. and let* $r^2(s,t) = E(\eta(t)-\eta(s))^2$. *Then for any* $\varepsilon > 0$ *and* $0 < \delta < 1$ *the following inequality holds:*

$$\begin{aligned} &P\{\sup_{\tau(s,t)<\delta}|\eta(s)-\eta(t)| > \varepsilon\} \\ &\le C\varepsilon^{-1}\left(\int_0^\delta (H([0,1],\tau,x))^{1/2}\,dx + \delta\ln^+\ln(4d\delta^{-1})^{1/2}\right), \end{aligned}$$

where $d = \sup_{s,t\in[0,1]}\tau(s,t)$ *and* $\ln^+ u = \ln(1\vee u)$.

In the case under consideration

$$\tau(s,t) = (E|Y(t) - Y(s)|^2)^{1/2} \le (F_1(t) - F_1(s))^{\alpha_1/2} \le C|t-s|^{\beta_1\alpha_1/2}.$$

After some rather rough estimates we obtain

$$I_{2,1} \le C\varepsilon^{-1}\delta^{\beta_1\alpha_1/4}.$$

To estimate $I_{1,1}$ one needs the following lemmas. The first two may be found in Billingsley (1968) and the third in Hahn (1978).

Lemma 4.14. *Let $[a,b] \subset [0,1]$. For $x \in D[0,1]$ the following inequality holds:*

$$\sup_{t\in[a,b]} |x(t)-x(a)| \le \sup_{t\in[a,b]} \min\{|x(t)-x(a)|,\ |x(b)-x(t)|\} + |x(b)-x(a)|.$$

Lemma 4.15. *Let $\xi_1,\dots,\xi_m$ be r.v.'s and let $S_0 = 0$ and $S_k = \sum_{i=1}^k \xi_i$, $k = 1,2,\dots,m$. Suppose that there exist non-negative numbers $u_1,\dots,u_m$ such that for any $\lambda > 0$, $0 \le i \le j \le k \le m$ the inequality*

$$P\{|S_j - S_i| \ge \lambda,\ |S_k - S_j| \ge \lambda\} \le \lambda^{-\gamma}\left(\sum_{n=i}^k u_n\right)^\alpha$$

holds for some $\gamma \ge 0$ and $\alpha > 1$. Then for all $\lambda > 0$

$$P\{\max_{0\le j\le m} \min\{|S_j|,\ |S_m - S_j|\} \ge \lambda\} \le K_{\gamma,\alpha}\lambda^{-\gamma}\left(\sum_{i=1}^m u_i\right)^\alpha,$$

where $K_{\gamma,\alpha}$ is a constant depending on γ and α only.

Lemma 4.16. *Suppose that X satisfies the conditions (4.8) and (4.9). Then for all $\lambda > 0$, $0 \le s \le t \le u \le 1$*

$$\lambda^4 P\{|S_n(t) - S_n(s)| \ge \lambda,\ |S_n(u) - S_n(t)| > \lambda\} \le (G(u) - G(s))^{\mu_1},$$

where $\mu_1 = \alpha_2 \wedge 2\alpha_1 > 1$ and $G(u) = 2^{1/\mu_1}(n^{-1/\mu_1}F_2(u) + 3^{1/\mu_1}F_1(u))$.

Applying Lemmas 4.14–4.16, we deduce that

$$I_{1,1} \le \sum_{i=1}^{N-1} (P\{R_i > \varepsilon/2\} + P\{|S_n(t_{i+1}) - S_n(t_i)| > \varepsilon/2\}), \tag{4.19}$$

$$P\{R_i > \varepsilon/2\} \le C\varepsilon^{-4}(G(t_{i+1}) - G(t_i))^{\mu_1},$$

where $R_i = \sup_{t_i\le t\le t_{i+1}} \min\{|S_n(t) - S_n(t_i)|,\ |S_n(t_{i+1}) - S_n(t)|\}$. To estimate the second term on the right-hand side of (4.19) we use (4.15) with $k = 1$ and obtain

$$P\{|S_n(t_{i+1}) - S_n(t_i)| > \varepsilon/2\} \le C\varepsilon^{-4}E(S_n(t_{i+1}) - S_n(t_i))^4$$
$$\le C\varepsilon^{-4}\big(n^{-1}(F_3(t_{i+1}) - F_3(t_i))^{\alpha_3} + 3(F_1(t_{i+1} - (F_1(t_i))^{2\alpha_1}\big).$$

Finally, applying (4.16) for all F_i, $i = 1, 2, 3$, we have

$$I_{1,1} \le C\varepsilon^{-4}\big((\delta^{\beta_2}n^{-1/\mu_1} + \delta^{\beta_1})^{\mu_1 - 1} + \delta^{\beta_3(\alpha_3-1)}n^{-1}\big).$$

In order to apply Theorem 4.5 to estimate $\Delta_n^N(\lambda)$ and $\Delta_n^N(\lambda - \varepsilon)$ we can use the estimate for I_2, obtained above to bound $D_n(\delta)$. It is not difficult to deduce that

$$D_n(\delta) \le C\max(n^{-1/6},\ n^{1/12}\delta^{\beta_1\alpha_1/8}).$$

It remains to collect all estimates which we have for the quantities entering inequality (4.18) and to choose the parameters ε and δ in an appropriate way. It is easy to see that if we put $\varepsilon \sim n^{-1/6}$ and $\delta \sim n^{-\rho}$ with some positive constant ρ depending on β_i, α_i, $i = 1, 2, 3$, and p, then all terms will not exceed $n^{-1/6}\lambda^{-3}$. It is worth mentioning that $\rho \to \infty$ in any of the following cases: if $\min_i \beta_i \to 0$, or if the α_i tend to their least possible values, or $p \to 3$. Also one can notice that under these stronger conditions which we had assumed for the final estimate of the remainder term there will be no term $\ln(1 + \lambda)$

§1.5. Rates of Convergence in Prokhorov and BL Metrics

Let $\mathcal{P}(B)$ denote the set of probability measures on B. We recall that the Prokhorov metric π and the bounded Lipschitz (BL) metric ϱ_{BL} on $\mathcal{P}(B)$ are defined as follows: for $\mu, \nu \in \mathcal{P}(B)$

$$\pi(\mu,\nu) := \inf\{\varepsilon > 0 : \mu(F) \le \nu(F_\varepsilon) + \varepsilon \text{ for all closed sets } F \subset B\};$$

$$\varrho_{BL}(\mu,\nu) := \sup\left\{\left|\int_B f(x)(\mu - \nu)(dx)\right| : ||f||_{BL} \le 1\right\},$$

where F_ε denotes an ε-neighborhood of the set $F \subset B$ and

$$||f||_{BL} := \sup_{x\in B}|f(x)| + \sup_{x\ne y}|f(x) - f(y)|/||x - y||$$

for $f : B \to \mathbb{R}$. For the sake of brevity we shall write $\pi(\xi,\eta)$ and $\varrho_{BL}(\xi,\eta)$ instead of $\pi(\mathcal{L}(\xi),\mathcal{L}(\eta))$ and $\rho_{BL}(\mathcal{L}(\xi),\mathcal{L}(\eta))$, respectively. Recall that both metrics π and ϱ_{BL} metrize the weak convergence on $\mathcal{P}(B)$. Therefore the rate of convergence, estimated by means of these metrics, is of great interest. Results on the rate of convergence for the Prokhorov metric in finite-dimensional spaces may be found, for example, in Yurinskii (1977), Dehling (1983), Zaitsev (1987), and Bentkus (1984f). In the case of infinite-dimensional spaces we have the following negative result.

Theorem 5.1 (Senatov (1981)). *For any monotone sequence $b_n \downarrow 0$ there exist a Gaussian ℓ_2-r.e. Y and ℓ_2-r.e. X such that*

(i) $EX = EY = 0$, $cov\,X = cov\,Y$;
(ii) $P\{\|X\| \le 1\} = 1$;
(iii) $\liminf_{n\to\infty} \pi(S_n(X), Y)b_n^{-1} > 0$;
(iv) $\liminf_{n\to\infty} \varrho_{BL}(S_n(X), Y)b_n^{-1} > 0$.

Note that conditions (i), (ii) guarantee the CLT; therefore $\pi(S_n(X), Y) \to 0$ and $\varrho_{BL}(S_n(X), Y) \to 0$ as $n \to \infty$. Theorem 5.1 tells us that in order to have a convergence rate either in the Prokhorov metric or in the BL metric in infinite-dimensional spaces one needs stronger assumptions than the finiteness of the usual moments. The following results provide such conditions. Recall that the r.e.'s X, $Y \in B$ are such that $EX = EY = 0$, $cov\,X = cov\,Y$ and that Y is Gaussian.

Theorem 5.2. *Let r.e.'s X, $Y \in \ell_2$. Suppose that*

$$E\|X\|_\lambda^3 < \infty,$$

where $\|x\|_\lambda^2 = \sum_{i=1}^\infty (x_i/\lambda_i)^2$ and the sequence $\lambda = (\lambda_1, \lambda_2, \ldots)$, $\lambda_i > 0$, is such that

$$\sum_{i=1}^\infty \lambda_i^2 < \infty. \tag{5.1}$$

Then

$$\pi(S_n(X), Y) = O(n^{-1/8})$$

and

$$\varrho_{BL}(S_n(X), Y) = O(n^{-1/6}).$$

Theorem 5.3. *Let r.e.'s X, $Y \in c_0$. Suppose that*

$$E\|X\|_2^3 < \infty,$$

where $\|x\|_2^2 = \sum_{i=1}^\infty x_i^2$. Then

$$\pi(S_n(X), Y) = O(n^{-1/8})$$

and

$$\varrho_{BL}(S_n(X), Y) = O(n^{-1/6}).$$

Therefore the finiteness of the third moment under an appropriate semi-norm, stronger than the original one, is sufficient to derive a convergence rate in the Prokhorov and BL metrics. Theorem 5.2 is due to Bentkus and Račkauskas (1984). Theorem 5.3 was proved by Bentkus (the proof is contained in Paulauskas and Račkauskas (1989)) and, in fact, is stronger than Theorem 5.2. Now we shall state more general results than Theorem 5.2. To this end, let a Hilbert space $H \subset B$ be linearly and continuously imbedded in the Banach space B (therefore, without loss of generality, one may assume that $\|x\|_H \ge \|x\|_B$ for all $x \in H$). The imbedding operator $H \hookrightarrow B$ is said

to be γ-radonifying if the Gaussian measure ν of the cylinder sets in B with characteristic functional

$$\int_B \exp(ix(y))\nu(dy) = \exp(-\|x\|_H^2/2)$$

for $x \in B^* \subset H^* = H$ is a σ-additive measure (see, e.g., Badrikian and Chevet (1974)). Put

$$\nu_{H,m} = \int_B \|x\|_H^m |\mathcal{L}(X) - \mathcal{L}(Y)|(dx).$$

Theorem 5.4 (Bentkus and Račkauskas (1984)). *Suppose that the imbedding $H \hookrightarrow B$ is γ-radonifying. Then there exists a finite constant C such that*

$$\pi(S_n(X), Y) \le C\nu_{H,3}^{1/4} n^{-1/8}; \tag{5.2}$$

$$\varrho_{BL}(S_n(X), Y) \le C \max(\nu_{H,3} n^{-1/2},\ \nu_{H,3}^{1/3} n^{-1/6}). \tag{5.3}$$

We note that the corresponding Hilbert space H for the space $B = \ell_2$ can be taken as

$$H = \ell_2(\lambda) := \left\{ x \in \ell_2 : |x|_H^2 := \sum_{i=1}^\infty (x_i/\lambda_i)^2 < \infty \right\},$$

where the sequence $\lambda = (\lambda_1, \lambda_2, \ldots)$ of positive numbers is such that (5.1) holds.

In Hilbert space, Yurinskii (1977) obtained

$$\pi(S_n(X), Y) = O(n^{-\alpha/(6+8\alpha)} \log n)$$

under the assumption of a finite third moment, where α is a positive parameter characterizing the behaviour of the eigenvalues of the covariance operator $cov\,Y$. Conditions, guaranteeing some logarithmic order of $\pi(S_n(X), Y)$ in the case X, $Y \in \ell_2$, may be found in Kukuš (1981, 1982). Lapinskas (1978) investigated the convergence rate in the CLT in the Prokhorov metric in Banach spaces with a Schauder basis. He imposed certain assumptions on the coordinates of X, Y and obtained the rate $O(n^{-1/21})$. For X, $Y \in c_0$ and $E\|X\|_1^3 < \infty$, where $\|\cdot\|_1$ is the ℓ_1-norm, a result of Senatov (1981) yields the estimate $\pi(S_n(X), Y) = O(n^{-1/8} \log^{3/4} n)$ (compare with Theorem 5.2).

Note that both estimates (5.2) and (5.3) are in general unimprovable. Appropriate examples are constructed by Bentkus (1987) (see also Senatov (1981) and Theorem 26 in Bentkus and Račkauskas (1984)). It should be noted that the condition $\nu_{H,3} < \infty$ is rather restrictive. For example, it can be shown that when the limiting Gaussian element Y is a Wiener process in $B = C[0,1]$, we automatically have $\nu_{H,3} = \infty$ if $\mathcal{L}(X) \ne \mathcal{L}(Y)$. Bentkus and Račkauskas (1984) proposed a method which enables one to obtain estimates under less restrictive assumptions. However, the order in n obtained by this method is somewhat worse. We shall not go into detail but confine ourselves to a particular example concerning the space $B = C[0,1]$. If ω is the modulus of continuity, then we denote by H_ω the space of all functions x on $[0,1]$ such that

$$||x||_\omega := \sup_{t\in[0,1]} |x(t)| + \sup_{t\neq s} |x(t)-x(s)|/\omega(|t-s|) < \infty.$$

In the case $\omega(s) = s^\alpha$, we write $||\cdot||_\alpha$, H_α instead of $||\cdot||_\omega$, H_ω, respectively. Theorems 5.5 and 5.6 below assume instead of $EX = EY = 0$, $cov\, X = cov\, Y$ that $EX(s) = EY(s) = 0$, $EX(t)X(s) = EY(t)Y(s)$ for all $s, t \in [0,1]$. The next result is a slightly improved version of Bentkus and Račkauskas (1984) (see Paulauskas and Račkauskas (1989)).

Theorem 5.5. *Let $X, Y \in C[0,1]$. Suppose that the following conditions are fulfilled:*

$$P\{X \in H_\alpha\} = P\{Y \in H_\alpha\} = 1$$

and

$$\nu_{\alpha,3} := \int_{C[0,1]} ||x||_\alpha^3 |\mathcal{L}(X) - \mathcal{L}(Y)|(dx) < \infty.$$

If $1/2 < \alpha \leq 1$, then there exists a constant $C(\alpha)$ such that

$$\pi(S_n(X), Y) \leq C(\alpha)\nu_{\alpha,3}^{1/4} n^{-1/8},$$

and

$$\varrho_{BL}(S_n(X), Y) \leq C(\alpha)\nu_{\alpha,3}^{1/3} n^{-1/6}.$$

If $0 < \alpha \leq 1/2$, then for each $\varepsilon > 0$ there exists a constant $C(\alpha, \varepsilon)$ such that

$$\pi(S_n(X), Y) \leq C(\alpha, \varepsilon) \max\big(\nu_{\alpha,3}^{1/3} n^{-1/6},\ (\nu_{\alpha,3} n^{-1/2})^{4\alpha/(9-2\alpha)-\varepsilon}\big), \tag{5.4}$$

and

$$\varrho_{BL}(S_n(X), Y) \leq C(\alpha, \varepsilon) \max\big(\nu_{\alpha,3}^{1/3} n^{-1/6}, (\nu_{\alpha,3} n^{-1/2})^{2\alpha/3-\varepsilon}\big). \tag{5.5}$$

In the case where Y is a Wiener process on $[0,1]$ and $E||X||_\alpha^3 < \infty$ with $\alpha = 1/2 + \delta$, $\delta < 0$, consequences of Theorem 5.5 are

$$\pi(S_n(X), Y) = o(n^{-1/8+\varepsilon}),$$

and

$$\varrho_{BL}(S_n(X), Y) = o(n^{-1/6+\varepsilon}),$$

where $\varepsilon = \varepsilon(\delta) \downarrow 0$ when $\delta \uparrow 0$.

Combining the method of Bentkus and Račkauskas (1984) with the results due to Bogachev (1988), one can strengthen (5.4) and (5.5) for $\alpha = 1/2$. Namely, the following result holds.

Theorem 5.6. *Let $X, Y \in C[0,1]$ and $\omega(t) = t^{1/2}(\log(1/t))^{-\beta}$, $\beta > 3/2$. Suppose that the following conditions are fulfilled:*

$$P\{X \in H_\omega\} = P\{Y \in H_\omega\} = 1$$

and

$$\nu_{\omega,3} := \int_{C[0,1]} ||x||_\omega^3 |\mathcal{L}(X) - \mathcal{L}(Y)|(dx) < \infty.$$

Then there exists a constant $C = C(\beta)$ such that

$$\pi(S_n(X), Y) \leq C\nu_{\omega,3}^{1/4} n^{-1/8};$$

and

$$\varrho_{BL}(S_n(X), Y) \leq C\nu_{\omega,3}^{1/3} n^{-1/6}.$$

The convergence rate in the CLT estimated by means of other distances than the Prokhorov metric or BL metric was considered by Zolotarev (1976a, b, 1977), Bentkus and Račkauskas (1984), Sakalauskas (1983), Liubinskas (1987), Rachev and Yukich (1989), and Rachev and Rüschendorf (1990). Bounds of the convergence rate in the infinite-dimensional invariance principle in the Prokhorov metric and other metrics were studied by Borovkov and Sakhanenko (1980), Borovkov (1984), Sakhanenko (1988) and Bentkus and Liubinskas (1989).

Chapter 2
Asymptotic Expansions

§2.1. Short Expansion

Throughout this chapter we shall use the notation $\varepsilon = \varepsilon(n) = n^{-1/2}$. Let $f : B \to F$ denote a function on a Banach space B such that the expectations $Ef(S_n)$, $Ef(Y)$ are well defined. We set

$$g(\varepsilon) = g(\varepsilon; f) = Ef(S_n),$$
$$g(0) = g(0; f) = Ef(Y)$$

We shall describe the general idea for constructing asymptotic expansions of the type

$$g(\varepsilon) = g(0) + a_1\varepsilon + \cdots + a_k\varepsilon^k + R \tag{1.1}$$

or, equivalently,

$$Ef(S_n) = Ef(Y) + a_1 n^{-1/2} + \cdots + a_k n^{-k/2} + R \tag{1.2}$$

(the so-called asymptotic power series). Here $a_1, \ldots, a_k$ are "known" coefficients and the remainder term $R = R_k(\varepsilon)$ usually satisfies $\varepsilon^{-k} R_k(\varepsilon) \to 0$ as $\varepsilon \to 0$. By choosing various functions f in (1.2), one can obtain asymptotic expansions for moments, probabilities, etc. We remark that the structure of our asymptotic expansions depends neither on the structure of f and Y nor on the dimension and the structure of Banach spaces B and F.

The following lemma is obvious.

Lemma 1.1. *If the remainder term $R = R_k(\varepsilon)$ satisfies $\varepsilon^{-k} R_k(\varepsilon) \to 0$ as $\varepsilon \to 0$ then the coefficients $a_1, \ldots, a_k$ in the expansion (1.1) are unique and for $s = 1, \ldots, k$*

$$a_s = \lim_{\varepsilon \to 0} \varepsilon^{-s}(g(\varepsilon) - g(0) - a_1\varepsilon - \cdots - a_{s-1}\varepsilon^{s-1}).$$

We shall demonstrate in detail the complete construction only in the case of the short asymptotic expansion

$$Ef(S_n) = Ef(Y) + a_1\varepsilon + R.$$

The general case differs from this only in cumbersome technical details and more complicated notation. Also we restrict ourselves to the case of functions $f : B \to F$ with bounded derivatives. This restriction is essential and considerably simplifies the estimation of the remainder term. However, we point out that, roughly speaking, to estimate the remainder term for non-smooth f, one has to apply additionally the methods developed for estimating the convergence rates in the CLT.

Theorem 1.2. *If $f : B \to F$ is a function of the class C_b^6 and $E||X||^4 < \infty$, then*

$$Ef(S_n) = Ef(Y) + \frac{1}{6}\varepsilon E f'''(Y)X^3 + R, \tag{1.3}$$

where

$$||R|| \le \varepsilon^2 C(f)\left(E||X||^4 + E||Y||^4 + (E||X||^3 + E||Y||^3)^2 + E||X||^3 E||Y||\right), \tag{1.4}$$
$$C(f) = ||f^{(4)}||_\infty + ||f^{(6)}||_\infty, \quad ||f||_\infty = \sup_{x \in B} ||f(x)||.$$

If B is a Hilbert space, then

$$||R|| \le C\varepsilon^2(||f^{(4)}||_\infty + ||f^{(6)}||_\infty)(1 + E||Y||^2)E||X||^4,$$

where C is an absolute constant.

The proof of Theorem 1.2 involves several steps (see Lemmas 1.3–1.6). To estimate the convergence rate, the following obvious algebraic identity is very useful (see §2.1):

$$\mu_1 \cdots \mu_n = \nu_1 \cdots \nu_n + \sum_{i=1}^n \mu_1 \cdots \mu_{i-1}(\mu_i - \nu_i)\nu_{i+1} \cdots \nu_n, \tag{1.5}$$

where $\mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_n$ are arbitrary measures and the multiplication is understood to be convolution of measures. For asymptotic expansions one needs to apply (1.5) iteratively several times. The number of iterations depends on the desired estimate of the remainder term. In the case of the short asymptotic expansions it is enough to use (1.5) twice, i.e. the identity

$$\mu_1 \cdots \mu_n = \nu_1 \cdots \nu_n + \sum_{i=1}^n \nu_1 \cdots \nu_{i-1}(\mu_i - \nu_i)\nu_{i+1} \cdots \nu_n + R, \tag{1.6}$$

where the remainder

$$R = \sum_{i=1}^n \sum_{j=1}^{i-1} \mu_1 \cdots \mu_{j-1}(\mu_j - \nu_j)\nu_{j+1} \cdots \nu_{i-1}(\mu_i - \nu_i)\nu_{i+1} \cdots \nu_n.$$

In the case of identically distributed summands, $\mu = \mu_1 = \cdots = \mu_n$, $\nu = \nu_1 = \cdots = \nu_n$ and (1.6) reduces to

$$\mu^n = \nu^n + n\nu^{n-1}(\mu - \nu) + R, \tag{1.7}$$

with

$$R = \sum_{i=1}^{n-1}(n - i)\mu^{i-1}(\mu - \nu)^2\nu^{n-i-1}.$$

Integrating f with respect to the measure μ^n, choosing $\mu = \mathcal{L}(\varepsilon X)$, $\nu = \mathcal{L}(\varepsilon Y)$ and applying (1.7), we obtain

$$Ef(S_n) = Ef(Y) + n\int_B Ef(\varepsilon\sqrt{n-1}\,Y + \varepsilon x)H(dx) + R, \tag{1.8}$$

where

$$H = \mathcal{L}(X) - \mathcal{L}(Y)$$

and

$$R = \sum_{i=1}^{n-1}(n - i)\int_B \int_B f(a_i + \varepsilon x + \varepsilon y)H(dx)H(dy),$$

and where r.e. a_i has the distribution of $\varepsilon X_1 + \cdots + \varepsilon X_{i-1} + \varepsilon\sqrt{n-i-1}\,Y$.

Lemma 1.3. *For all $a \in B$ and $\varepsilon > 0$ the following estimates hold:*

$$\left\|\int_B f(a + \varepsilon x)H(dx) - \frac{1}{6}\varepsilon^3 \int_B f'''(a)x^3 H(dx)\right\|$$
$$\le \frac{1}{24}\varepsilon^4 ||f^{(4)}||_\infty (E||X||^4 + E||Y||^4), \tag{1.9}$$
$$\left\|\int_B \int_B f(a + \varepsilon x + \varepsilon y)H(dx)H(dy)\right\|$$
$$\le \frac{1}{36}\varepsilon^6 ||f^{(6)}||_\infty (E||X||^3 + E||Y||^3)^2. \tag{1.10}$$

Proof. To prove (1.9) it is sufficient to apply Taylor's formula (see, e.g., Cartan (1971))

$$h(u+v)=\sum_{j=0}^{s}h^{(j)}(u)v^j/j!+\int_0^1(1-\tau)^s h^{(s+1)}(u+\tau v)v^{s+1}d\tau/s!$$

with $h=f$, $u=a$, $v=\varepsilon x$, $s=3$ and to note that the means and covariances of X and Y coincide. For the proof of (1.10) it is sufficient to apply Taylor's formula twice with $s=2$ and $v=\varepsilon x$ and $v=\varepsilon y$, respectively.

Lemma 1.4. *The following relation holds:*

$$Ef(S_n)=Ef(Y)+\varepsilon Ef'''(\varepsilon\sqrt{n-1}\,Y)X^3/6+R,$$

where

$$\|R\|\le\varepsilon^2(\|f^{(4)}\|_\infty+\|f^{(6)}\|_\infty)(E\|X\|^4+E\|Y\|^4+(E\|X\|^3+E\|Y\|^3)^2/3)/24.$$

Proof. The lemma is a consequence of the representation (1.8) and Lemma 1.3. We apply the estimate (1.9) to the integral in (1.8) with $a=\varepsilon\sqrt{n-1}\,Y$. The remainder term R in (1.8) is estimated with the help of (1.10) with $a=a_i$. Finally, we have

$$\int_B f'''(a)x^3H(dx)=Ef'''(a)X^3$$

due to the symmetry of Y.

Lemma 1.5.

$$\|Ef'''(Y)X^3-Ef'''(\varepsilon\sqrt{n-1}\,Y)X^3\|\le\varepsilon\|f^{(4)}\|_\infty E\|X\|^3\|E\|Y\|.$$

Proof. The r.e. Y has the same distribution as $\varepsilon\sqrt{n-1}\,Y+\varepsilon Y_1$, where Y_1 is an independent copy of Y. Therefore it is sufficient to apply Taylor's formula to $Ef'''(Y)X^3$ with $h=f'''$, $u=\varepsilon\sqrt{n-1}\,Y$, $v=\varepsilon Y_1$ and $s=0$.

Lemma 1.6. *If B is a Hilbert space, then $E\|X\|^2=E\|Y\|^2$. Furthermore $E\|Y\|^p\le C(p)E\|X\|^p$, $p>0$.*

Proof. The equality $E\|X\|^2=E\|Y\|^2$ is obvious since $\|x\|^2$ in a Hilbert space is a quadratic form and the covariances of X and Y coincide. Therefore the well-known inequality $E\|Y\|^p\le C(p)(E\|Y\|^2)^{p/2}$ and the Hölder inequality help to complete the proof of the lemma.

§2.2. The Smooth Case

In this section we shall consider the results obtained without (explicit) conditions similar to the classical Cramér condition

$$\limsup_{|t|\to\infty}|E\exp(itX)|<1$$

for the characteristic function of a r.v. $X\in\mathbb{R}$. In the "smooth case" the formulations of results are not that overloaded by conditions and technical details. Therefore we shall consider different forms of expansions and certain technical aspects. We begin with the description of the coefficients $a_1,\dots,a_k$ in the general asymptotic expansion. We shall show that

$$a_s=EP_sf(Y),\quad s=0,1,2,\dots,\tag{2.1}$$

where the P_s are certain random differential operators. Let us repeat once more that the construction of P_s is universal and does not depend on the specific structure of X, Y, f or the spaces B, F. The coefficients a_s have the form (2.1) even when f is not differentiable, for example, when $f(x)=\chi_A(x)$ is the indicator function of a set $A\subset B$. The only difference in the case of a non-smooth function f is that formula (2.1) has to be interpreted in an appropriate way (see §2.3).

The Edgeworth–Cramér polynomials $E_k=E_k(m_2,\dots,m_{k+2})$ of formal commuting variables $m_2,\dots,m_{k+2}$ (the so-called "moment" variables) are determined as the coefficients in the formal power series expansion

$$\exp\left(t^{-2}\left[\ln\left(1+\sum_{k=2}^{\infty}m_kt^k/k!\right)-m_2t^2/2\right]\right)=\sum_{k=0}^{\infty}E_k(m_2,\dots,m_{k+2})t^k.$$

For instance,

$$\begin{aligned}E_0&=1,\ E_1=m_3/6,\\ E_2&=-m_2m_2/8+m_3m_3/72+m_4/24,\\ E_3&=-m_2m_3/12-m_2m_2m_3/48-m_3m_3m_3/1296+m_3m_4/144+m_5/120.\end{aligned}$$

Write

$$E_k=\sum a_k(i_1,\dots,i_s)m_{i_1}\cdots m_{i_s},$$

where the sum is taken over all integers $i_1,\dots,i_s$ such that $2\le i_1\le\dots\le i_s$. Clearly, only a finite number of coefficients $a_k(\dots)$ are non-zero. Define the polynomials

$$P_k=P_k(z_1,\dots,z_k)=\sum a_k(i_1,\dots,i_s)z_1^{i_1}\cdots z_s^{i_s}\tag{2.2}$$

of commuting variables $z_1,\dots,z_k$. For instance,

$$P_2 = -\frac{1}{8}z_1^2z_2^2 + \frac{1}{72}z_1^3z_2^3 + \frac{1}{24}z_1^4.$$

If $h \in B$, then one may introduce the differential operator $D(h)$ as follows:

$$D(h)f(x) = f'(x)h = \lim_{t\to 0} t^{-1}(f(x+th) - f(x))$$

(the so-called directional derivative). The differential operators $D(h)$, $h \in B$, and $1 \in \mathbb{R}$, generate a natural commutative algebra over the field of real numbers. Therefore one can define the random differential operators

$$P_k = P_k(D(X_1), \dots, D(X_k))$$

and coefficients $a_s = EP_sf(Y)$. For instance,

$$P_2 = -\frac{1}{8}D^2(X_1)D^2(X_2) + \frac{1}{72}D^3(X_1)D^3(X_2) + \frac{1}{24}D^4(X_1).$$

Lemma 2.1. *Let $E||X||^{k+2} < \infty$ and let function $f \in C^{3k}(B;F)$. Then there exists an $\alpha = \alpha(\mathcal{L}(Y)) > 0$ such that the coefficients*

$$a_s = EP_sf(Y), \ s = 0, 1, \dots, k,$$

are well defined when

$$\max_{0\le i\le 3k} \sup_{x\in B} \exp(-\alpha||x||^2)||f^{(i)}(x)|| < \infty. \tag{2.3}$$

Since we want to avoid complications connected with measurability we assume separability of B and F. The condition $E||X||^{k+2} < \infty$ allows one to interpret expectations in the sense of Bochner instead of more complicated definitions. Condition (2.3) is connected with the well-known integrability properties of the norm of a Gaussian r.e. (the Skorohod – Fernique – Landau – Shepp Theorem). The condition $f \in C^{3k}$ is unnecessarily strong and will be weakened later on.

Theorem 2.2. *Suppose that a r.e. $X \in B$ satisfies the CLT. Let $E||X||^{k+3} < \infty$, function $f \in C^{3k+3}(B;F)$ and*

$$\sup_{x\in B}(1+||x||)^{-k-3}||f(x)|| < \infty. \tag{2.4}$$

Then there exists $\alpha = \alpha(\mathcal{L}(Y)) > 0$ such that the asymptotic expansion

$$Ef(S_n) = Ef(Y) + \sum_{s=1}^{k} n^{-s/2}EP_sf(Y) + R \tag{2.5}$$

is well defined and $||R|| = O(n^{-(k+1)/2})$ provided

$$\sup_{x\in B} \exp(-\alpha||x||_B)||f^{(i)}(x)|| < \infty \tag{2.6}$$

for all $i = 1, \dots, 3k+3$.

We now make some comments about Theorem 2.2 and provide some references. We are not concerned with the finite-dimensional case (see, e.g., Petrov (1975), Bhattacharya and Rao (1976), Götze and Hipp (1979), etc.). We mention here only that the formal power series expansion with Edgeworth–Cramér polynomials is contained, e.g., in Bikelis (1973). The definition of the terms $EP_kf(Y)$ of the asymptotic expansions via random differential operators P_k may be found in Götze (1981). The expansion (2.5) in the case $k = 0$ (i.e., the estimate of the convergence rate) under the condition $f \in C_b^3(B;\mathbb{R})$ was obtained by Paulauskas (1976b) and Zolotarev (1976b). Götze (1981) proved (2.5) under the assumption $f \in C_b^{3k+3}(B;\mathbb{R})$. Bentkus (1984b) proved that the short asymptotic expansion (2.5) is valid under the weaker moment and differentiability conditions $E||X||^3 < \infty$, $f \in C_b^3(H;\mathbb{R})$ (H is a Hilbert space), while the estimate of the remainder term becomes worse: $R = o(n^{-1/2})$ instead of $R = O(n^{-1})$. Theorem 2.2 in the case of a Hilbert space is contained in Bentkus (1984d). In the general case Theorem 2.2 follows from more general and more precise results due to Bentkus (1986d). It should be remarked also that all papers, just mentioned, contain more or less explicit estimates of the remainder term.

The construction of asymptotic expansions is usually based on the formula obtained by iterating (1.6), that is,

$$\mu_1 \cdots \mu_n = \nu_1 \cdots \nu_n + A_1 + \dots + A_k + R, \tag{2.7}$$

where $\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_n$ are arbitrary measures and

$$A_s = \sum_{card\, \alpha = s} \prod_{i\in\alpha}(\mu_i - \nu_i) \prod_{j\notin\alpha} \nu_j. \tag{2.8}$$

The summation is over all possible sets $\alpha = \{i_1, \dots, i_s\}$ of integers $1 \le i_s < i_{s-1} < \dots < i_1 \le n$ and the second product is over all integers $j \notin \alpha$ such that $1 \le j \le n$;

$$R = \sum_{card\, \alpha = k+1} \prod_{i} \mu_i \prod_{j\in\alpha}(\mu_j - \nu_j) \prod_{l\notin\alpha} \nu_l,$$

where the summation is similar to that in (2.8), the first product is taken over all integers $i \ge 1$, $i < i_{k+1}$ and the third product is over all integers $l \notin \alpha$ such that $i_{k+1} < l \le n$. In the identically distributed case $\mu = \mu_1 = \dots = \mu_n$, $\nu = \nu_1 = \dots = \nu_n$, the formulae for A_s and R reduce to

$$A_s = \binom{n}{s}\nu^{n-s}(\mu-\nu)^s, \quad \binom{n}{s} = n!/[s!(n-s)!],$$

and

$$R = \sum_{i=1}^{n-k}\binom{n-i}{k}\mu^i(\mu-\nu)^{k+1}\nu^{n-i-k-1}.$$

In the i.i.d. case, integrating (2.7) we obtain

$$Ef(S_n) = Ef(Y) + \sum_{s=1}^{k} \binom{n}{s} \varphi_s(\varepsilon) + R, \tag{2.9}$$

where

$$\varphi_s(\varepsilon) = E\int_B \cdots \int_B f(\varepsilon W_s + \varepsilon x_1 + \cdots + \varepsilon x_s) H(dx_1) \cdots H(dx_s),$$

$$R = \sum_{i=1}^{n-k} \binom{n}{s} R_i,$$

$$R_i = E\int_B \cdots \int_B f(\varepsilon T + \varepsilon x_1 + \cdots + \varepsilon x_{k+1}) H(dx_1) \cdots H(dx_{k+1}),$$

and the r.e.'s W_s and T have the following distributions:

$$\mathcal{L}(W_s) = \mathcal{L}(\sqrt{n-s}\,Y) = \mathcal{L}(Y_1 + \cdots + Y_{n-s}),$$
$$\mathcal{L}(T) = \mathcal{L}(X_1 + \cdots + X_i + Y_1 + \cdots + Y_{n-i-k-1}).$$

It seems that expansions of type (2.7), (2.8) were explicitly published for the first time by Bergström (1951). In any case, they are traditionally called "Bergström's expansions".

Let us discuss the method and results presented in the paper of Götze (1981). The main idea of his paper is that to construct the asymptotic expansion for $Ef(S_n)$, one does not need the differentiability of f. It is sufficient to have the differentiability of the functions φ_s and R_i in (2.9). The gain is based on the interchanging of the order of integration and differentiation and on the property of integration as a "smoothing operation". Put

$$U(\eta_1, \ldots, \eta_q) = Ef(Y + \eta_1 X_1 + \cdots \eta_q X_q),$$

$$U_i(\varepsilon_1, \ldots, \varepsilon_q | \eta_1, \ldots, \eta_q) = Ef(\varepsilon W_i + \varepsilon_1 Y_1 + \cdots + \varepsilon_q Y_q + \eta_1 X_1 + \ldots \eta_q X_q),$$

where the r.e. W_i satisfies

$$\mathcal{L}(W_i) = \mathcal{L}(X_1 + \cdots + X_i + Y_{i+1} + \cdots + Y_n),\ 1 \le i \le n.$$

Theorem 2.3. (Götze (1981)). *Suppose that the functions U_i are differentiable. Let*

$$C_n := \sup \left| \left(\frac{\partial}{\partial \varepsilon_1}\right)^i \left(\frac{\partial}{\partial \eta_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial \eta_q}\right)^{j_q} U_i(\varepsilon_1, 0, \ldots, 0|0, \eta_2, \ldots, \eta_q) \right|,$$

where sup *is taken over all $1 \le i \le n$, all $0 \le \varepsilon_1, \eta_2, \ldots, \eta_q \le n^{-1/2}$ and all $i, j_1, \ldots, j_q \le k+3$ such that $k+3 \le i+j_1+\ldots+j_q \le 3(k+1)$. Furthermore, suppose that*

$$\left(\frac{\partial}{\partial \varepsilon_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial \varepsilon_q}\right)^{i_q} \left(\frac{\partial}{\partial \eta_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial \eta_q}\right)^{j_q} U_i(0, \ldots, 0|0, \ldots, 0)$$
$$= \left(\frac{\partial}{\partial \eta_1}\right)^{i_1+j_1} \cdots \left(\frac{\partial}{\partial \eta_q}\right)^{i_q+j_q} U_i(0, \ldots, 0|0, \ldots, 0)$$

when $i_1 j_1 = \cdots = i_q j_q = 0$ for all $1 \le i \le n$, all $i_1, \ldots, i_q \le 2$ such that $i_1 + j_1 \le k+3, \ldots, i_q + j_q \le k+3$ and $i_1 + \cdots + i_q + j_1 + \cdots j_q \le 3(k+1)$. Then

$$\left| Ef(S_n) - \sum_{s=0}^{k} n^{-s/2} P_s U(0, \ldots, 0) \right| \le c(k) C_n n^{-(k+1)/2},$$

where the differentiable (non-random) *operator*

$$P_s = P_s\left(\frac{\partial}{\partial \eta_1}, \ldots, \frac{\partial}{\partial \eta_s}\right)$$

is defined via the polynomials P_s of (2.2).

It is possible to generalize Theorem 2.3 to apply in situations which are not directly connected with sums of independent r.e.'s. It occurs that the proof of this result is not probability-theoretic at all. This was noted by Götze (1985). Suppose that a sequence h_n, $n = 1, 2, \ldots$, of numbers is given, for example, $h_n = Ef(S_n)$, or the h_n are probabilities related to n observations, etc. Furthermore, suppose that it is possible to introduce "weights" $\varepsilon_1, \ldots, \varepsilon_n$ for the observations and to determine a sequence $h_n(\varepsilon_1, \ldots, \varepsilon_n)$, $0 \le \varepsilon_1, \ldots, \varepsilon_n \le n^{-1/2}$, of functions such that

$$h_n = h_n(n^{-1/2}, \ldots, n^{-1/2}) + O(n^{-(k+1)/2}).$$

For example, if $h_n = Ef(S_n)$, it is natural to set

$$h_n(\varepsilon_1, \ldots, \varepsilon_n) = Ef(\varepsilon_1 X_1 + \cdots + \varepsilon_n X_n).$$

Götze (1985) has shown under certain natural conditions (in particular, differentiability conditions similar to that in Theorem 2.3 in case $k = 0$) that the "central limit theorem" holds. That is for fixed $\eta_1, \ldots, \eta_r$ there exist limits

$$h_\infty = \lim_{n \to \infty} h_n(n^{-1/2}, \ldots, n^{-1/2}),$$
$$h_\infty(\eta_1, \ldots, \eta_r) = \lim_{m \to \infty} h_{m+r}(\eta_1, \ldots, \eta_r, n^{-1/2}, \ldots, n^{-1/2}).$$

Using the functions $h_n(\varepsilon_1, \ldots, \varepsilon_n)$ and $h_\infty(\eta_1, \ldots, \eta_r)$, one can define functions similar to the functions U and U_i in Theorem 2.3. Under differentiability conditions similar to those in Theorem 2.3, this allows one to obtain an asymptotic expansion of $h_n(n^{-1/2}, \ldots, n^{-1/2})$ analogous to that in Theorem 2.3. For further details in this direction and some applications to mathematical statistics, see Götze (1985).

Theorem 2.3 is valid for non-differentiable f, too. It has been proved without an explicit condition similar to the classical Cramér condition which is replaced here by differentiability conditions. In applications the verification of these conditions is the difficult task.

Asymptotic expansions with the Edgeworth–Cramér polynomials have the classical structure of asymptotic power series. A disadvantage of these expansions is that their existence is guaranteed only when X has all moments up to the order $k+2$. Furthermore, the structure of these expansions is not simple and convenient in all cases. For instance, the obvious fact that $EP_k f(Y) = 0$ when $\mathcal{L}(X) = \mathcal{L}(Y)$ needs a special proof when using the definitions of the P_k only. Most general and simple are the Bergström expansions (2.7), (2.9). Unfortunately, these expansions are not too informative. Bentkus (1984d) introduced certain expansions of intermediate type

$$Ef(S_n) = Ef(Y) + a_1(\varepsilon) + \cdots + a_k(\varepsilon) + R, \qquad (2.10)$$

where the functions $a_s(\varepsilon)$ satisfy

$$|a_s(\varepsilon)| \le C_s \varepsilon^s$$

with some constants $C_s < \infty$. These expansions have the following properties. They exist more frequently than Edgeworth expansions. They arise in a natural way from Bergström expansions. The closeness of $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ is taken into account and therefore it is possible to pass to the case of stable Y. They can be used to derive Edgeworth expansions and they lead to more general and exact results. But a drawback of these expansions is that the terms $a_s(\varepsilon)$ are not uniquely determined in contrast to the power series expansion. In order to describe the construction of $a_s(\varepsilon)$, consider the random differential operator

$$Q_i = Q_i(l^{(i)}) = (D^{l_1}(X_1) - D^{l_1}(Y_1)) \cdots (D^{l_i}(X_i) - D^{l_i}(Y_i)),$$

where $l^{(i)} = (l_1, \ldots, l_i)$ is a non-negative multi-index. If a mapping f has sufficiently many bounded derivatives, then one can define

$$a_s(\varepsilon) = \sum_{i=1}^{s} \sum_{|l^{(i)}|=2i+s} \left(n^{-i} \binom{n}{i} / l^{(i)}! \right) n^{-s/2} EQ_i f(\tau Y),$$

where $|l^{(i)}| = l_1 + \cdots l_i$, $\tau = (1 - i/n)^{1/2}$ and the second sum is over the $l^{(i)}$ satisfying $l_1 \ge 3, \ldots, l_i \ge 3$. Here $l^{(i)}! = l_1! \cdots l_i!$ and $\binom{n}{i} = n!/[i!(n-i)!]$.

Put $X^t = X\chi\{||X|| \le \sqrt{n}\}$, $X_t = X - X^t$,

$$L_p = n^{-(p-2)/2} E||X^t||^p, \quad \Lambda_p = n^{-(p-2)/2} E||X_t||^p.$$

Note that

$$\Lambda_{k+2} + L_{k+3} \le \varepsilon^{k+1} E||X||^{k+3}.$$

Lemma 2.4. *Under the conditions of Lemma 2.1,*

$$||a_s(\varepsilon)|| \le C\varepsilon^s E||X||^{s+2},$$

$$\left\| \sum_{s=1}^{k} [a_s(\varepsilon) - \varepsilon^s EP_s f(Y)] \right\| \le C(\Lambda_{k+2} + L_{k+3}),$$

where the constant $C = C(k, \mathcal{L}(Y))$.

Sometimes it is very useful to apply a truncation procedure, i.e., to replace $Ef(S_n)$ by $Ef(S_n^t)$, where $S_n^t = n^{-1/2}(X_1^t + \cdots X_n^t)$. Usually the difference $Ef(S_n) - Ef(S_n^t)$ can be easily estimated. As an example we state the following almost obvious Lemma 2.5 for bounded f (see Sazonov and Zalesskii (1985), and Bentkus (1986d) for the results concerning the case of unbounded f).

Lemma 2.5. *The following estimate holds:*

$$||Ef(S_n) - Ef(S_n^t)|| \le 2n||f||_\infty P\{||X|| \ge \sqrt{n}\}.$$

Lemma 2.5 (and its generalizations) allows one to replace the analysis of asymptotic expansions for $Ef(S_n)$ by the analysis of the asymptotic expansion

$$Ef(S_n^t) = Ef(Y) + a_1^t(\varepsilon) + \cdots + a_k^t(\varepsilon) + R, \qquad (2.11)$$

where the $a_s^t(\varepsilon)$ are determined by replacing $X_1, X_2, \ldots$ in the definition of $a_s(\varepsilon)$ by $X_1^t, X_2^t, \ldots$.

Lemma 2.6 (Compare with Lemma 2.1). *There exists $\alpha = \alpha(\mathcal{L}(Y)) > 0$ such that the condition*

$$\sup_{x\in B} \exp(-\alpha||x||^2)||f^{(i)}(x)|| < \infty, \ 0 \le i \le 3k,$$

ensures the existence of the asymptotic expansion (2.11). Moreover,

$$||Ef(Y)|| \le C,$$
$$||a_s^t(\varepsilon)|| \le C(\Lambda_2 + L_{s+2}), \ 1 \le s \le k,$$

where $C = C(k, \mathcal{L}(Y))$.

Theorem 2.7 (Compare with Theorem 2.2). *Suppose that a r.e. X satisfies the CLT and that $E||X||^2 < \infty$. Then there exists $\alpha = \alpha(\mathcal{L}(Y)) > 0$ such that the condition*

$$\sup_{x\in B} \exp(-\alpha||x||)||f^{(i)}(x)|| < \infty, \ 0 \le i \le 3k+3,$$

ensures the estimate

$$||R|| \le C(\mathcal{L}(X), k)(\Lambda_2 + L_{k+3})$$

of the remainder term in (2.11). If in addition $E||X||^{k+2} < \infty$, then

$$||a_s^t(\varepsilon) - a_s(\varepsilon)|| \le C\Lambda_{s+2}, \ 1 \le s \le k.$$

Further details and results concerning expansions (2.10), (2.11) may be found in Bentkus (1984d, 1986d).

We say that a Banach space B is of the class C_b^s (briefly $B \in C_b^s$) if the norm function $g(x) = ||x||$ is s-times continuously Fréchet differentiable on the open set $B - \{0\}$ and $\sup\{||g^{(i)}(x)|| : ||x|| = 1\} < \infty$, $1 \le i \le s$. In the next theorem concerning expansions for moments we assume that a r.e. X (or Y) is not finite-dimensional. This means that $P\{X \in E\} = 0$ if $E \subset B$ is a finite-dimensional subspace of B. This assumption does not restrict the generality since otherwise we may apply finite-dimensional results.

Theorem 2.8. *Suppose that a Banach space $B \in C_b^{3k+3}$. Then the asymptotic expansion*

$$Ef(S_n) = Ef(Y) + \sum_{s=1}^{k} \varepsilon^s EP_s f(Y) + R$$

is well defined and the remainder term satisfies $R = O(n^{-(k+1)/2})$ in the following cases:

(i) $f(x) = ||x||^p$ *for some* $p > 0$, $p < k+3$, *and* $E||X||^{k+3} < \infty$;
(ii) $f(x) = ||x||^p$ *for some* $p \ge k+3$ *and* $Ef(X) < \infty$;
(iii) $f(x) = \exp(a||x||^\alpha)$ *for some* $a > 0$, $0 < \alpha < 1$, *and* $Ef(X) < \infty$.

The first estimate of the convergence rate for the moments in an infinite-dimensional Hilbert space was given by Zalesskii and Sazonov (1984). Their proof includes integration of a nonuniform estimate for the probabilities and is based on the use of characteristic functions. Rhee and Talagrand (1984) constructed an example showing that even in a Banach space with very smooth norm, moment conditions alone cannot guarantee any convergence rate for the probabilities. Hence it turns out that the method of Zalesskii and Sazonov (1984) cannot be generalized directly to the case of Banach spaces. Using another method Bentkus (1984g, 1986d) generalized the result of Sazonov and Zalesskii to the case of a Banach space with sufficiently smooth norm and removed certain unnecessary restrictions. The asymptotic expansion of Theorem 2.8 is a consequence of more general and more exact results due to Bentkus (1986d).

§2.3. Asymptotic Expansions for Probabilities

Let A be a subset of a Banach space B. This section is devoted to asymptotic expansions for the probability $P\{S_n \in A\}$. If we choose $f(x) = \chi_A(x)$, then in the i.i.d. case we can rewrite the short Bergström expansion (1.8) in the following form

$$\begin{aligned} P\{S_n \in A\} &= P\{Y \in A\} + n \int_B Ef(Z + \varepsilon x) H(dx) + R \\ &= \nu(A) + n \int_B \Phi(A - \varepsilon x) H(dx) + R, \end{aligned} \tag{3.1}$$

with Gaussian measure $\Phi = \mathcal{L}(Z)$, $Z = \varepsilon\sqrt{n-1}\,Y$, and $H = \mathcal{L}(X) - \mathcal{L}(Y)$. The construction of asymptotic expansions in the case of smooth f was based on transforming $nE\int_B f(Z+\varepsilon x)H(dx)$ into $\varepsilon ED^3(X)f(Y)/6$ (if we leave the problem of estimating R aside). This, in turn, was based on expanding the smooth function $x \to f(Z+\varepsilon x)$ in a Taylor series in powers of εx. If f is non-differentiable, then such direct expansion is impossible. But assuming that the function

$$x \to \Phi(A - \varepsilon x) \tag{3.2}$$

is sufficiently smooth, we may repeat this argument with respect to the integral on the right-hand side of (3.1). Differentiability of the function (3.2) naturally leads to the notion of differentiable measure (see Averbuch, Smolyanov and Fomin (1971) and Daletskii and Fomin (1983) for more information concerning differentiable measures).

Suppose that $\Phi : \mathcal{A} \to \mathbb{R}$ is a set function (not necessarily additive or σ-additive) defined on a class $\mathcal{A}$ of subsets $A \subset B$. Let us suppose also that $\mathcal{A}$ is invariant under translations, $A \in \mathcal{A}$, $h \in B \Rightarrow A + h \in \mathcal{A}$. Then we can define (if it exists) the (directional) derivative

$$D(h)\Phi(A) = \lim_{t\to 0}\{\Phi(A - th) - \Phi(A)\}/t. \tag{3.3}$$

We mention that the traditional definition would require $\Phi(A + th)$ in (3.3). We did choose $\Phi(A - th)$ to keep the same notation for asymptotic expansions as in the previous sections. The first derivative $D(h)\Phi : \mathcal{A} \to \mathbb{R}$ is a set function. Hence we can define the successive derivatives $D(h_1)\cdots D(h_s)\Phi$ iteratively. This allows one to well define such quantities as, e.g., $EP_s\nu(A)$, where $P_s = P_s(D(X_1), \dots, D(X_s))$ is the random differential operator determined via the Edgeworth–Cramér polynomials. In the same way all formulas from the previous section can be correctly interpreted for the case of the non-smooth function $f(x) = \chi_A(x)$.

Concerning asymptotic expansions in the infinite-dimensional case, it is generally accepted to define the set A through a function $F : B \to \mathbb{R}$, i.e., to let $A = A_{r,F} = \{x \in B : F(x) < r\}$. The results obtained so far concern cases where F is sufficiently smooth or $F(x) = ||x||^p$, $p > 0$, in a Banach space with sufficiently smooth norm. Let us start with the case $F(x) = ||x||$, the norm of Hilbert space H. Let $V_{r,a} = \{x \in H : ||x - a|| < r\}$.

Theorem 3.1. *Suppose that r.e. $X \in H$ and is not finite-dimensional. If $E||X||^p < \infty$ for some $2 \le p < 3$, then for all $a \in H$, $r \in \mathbb{R}$*

$$P\{S_n \in V_{r,a}\} = P\{Y \in V_{r,a}\} + R \tag{3.4}$$

with $R = o(n^{-(p-2)/2})$. *If* $E||X||^p < \infty$ *for some* $3 \le p < 4$, *then*

$$P\{S_n \in V_{r,a}\} = P\{Y \in V_{r,a}\} + \frac{1}{6}\varepsilon ED^3(X)\nu(V_{r,a}) + R, \tag{3.5}$$

with $R = o(n^{-(p-2)/2})$. *Furthermore,* $ED^3(X)\nu(V_{r,a}) = 0$ *if* $a = 0$ *or if* X *is symmetric.*

We note that the theorem is valid without conditions like Cramér's condition for the c.f. The paper by Bentkus and Zalesskii (1985) contains an example with an infinite-dimensional $X \in H$ which shows that the distribution function $r \to P\{||S_n(X)|| < r\}$ has a jump bigger than C/n with some $C > 0$. Therefore the asymptotic expansions longer then those in Theorem 3.1 should contain additional discontinuous (with respect to r) terms or one should impose a condition similar to Cramér's condition. The following condition was introduced by Bentkus (1984e), and Nagaev and Chebotarev (1986).

The Cramér-type condition. There exists an operator $K \ge 0$ such that the operator $K \operatorname{cov} X$ is not finite-dimensional and

$$\limsup_{r\to\infty} \sup\{|E\exp(i(x,X))| : (Kx,x) = r\} < 1. \tag{3.6}$$

Remark 3.2. All of the results below hold if the range of the operator $K \operatorname{cov} X$ is of finite but sufficiently large dimension.

We note that one can never choose K to be the identity operator since it is easy to show that every r.e. X in an infinite-dimensional H satisfies (for each $r \ge 0$)

$$\sup\{|E\exp(i(x,X))| : ||x|| = r\} = 1.$$

Consider the condition

$$\int_{n^{3/4}\le|t|\le T} \frac{1}{|t|}|E\exp(it||S_n(X)||^2)|dt = O(1/T), \tag{3.7}$$

where T will be chosen later so that $1/T$ has the order of the desired error.

Usually one can verify (3.7) if Cramér's condition (3.6) is fulfilled (see Bentkus (1984e) and Nagaev and Chebotarev (1986)). Condition (3.7) is expressed in terms of the whole sum $S_n(X)$. Generally this is not accepted as a solution of the problem. Usually an estimate of the error is preferable when it is expressed in terms of one summand. We formulate the results while imposing (3.7) because it happens to be more convenient in applications (see, e.g., § 3.1 concerning ω^2 statistics); we are able to verify (3.7) but not the Cramér condition (3.6).

Theorem 3.3. *Let a r.e.* $X \in H$ *and let integer* $k \ge 2$. *Suppose that* $E||X||^{k+2} < \infty$. *Then the asymptotic expansion*

$$P\{S_n \in V_{r,a}\} = P\{Y \in V_{r,a}\} + \sum_{s=1}^{k} \varepsilon^s EP_s\nu(V_{r,a}) + R \tag{3.8}$$

is well defined. Here the P_s *are the random differential operators determined via Edgeworth–Cramér polynomials. Furthermore, suppose that condition (3.6) is fulfilled. If* $E||X||^{k+2+\alpha} < \infty$, *where* $0 \le \alpha \le 1$, *then*

$$R = o(n^{-(k+\alpha)/2})$$

and if $E||X||^{k+3} < \infty$, *then*

$$R = O(n^{-(k+1)/2}).$$

Moreover the result remains valid if instead of (3.6) condition (3.7) is fulfilled with any $T = o(n^{(k+\alpha)/2})$, $0 \le \alpha < 1$, *or* $T = O(n^{(k+1)/2})$, *respectively.*

The method of proving Theorem 3.1 and Theorem 3.3 was developed by Götze (1979, 1981, 1984), Yurinskii (1981, 1982), Zalesskii (1982), Nagaev (1983) and Bentkus (1984e). In these and later papers, various estimates of the remainder were obtained. Let us give a short review not concerning the first part of Theorem 3.1 (see § 1.1 for the convergence rates). The paper of Götze (1979) contains the very important symmetrization inequality (see § 1.1) which allows one to estimate the c.f. $E\exp(it||S_n + a||^2)$ for $|t| \le n^{1-\varepsilon}$, $\varepsilon > 0$. Also this paper gives bounds for the remainder term exact with respect to n but under somewhat restrictive moment conditions. Zalesskii (1982) proved Theorem 3.1 for $a = 0$. Theorem 3.1 and Theorem 3.3 follow from the results due to Bentkus (1984e). In that paper Theorem 3.1 and Theorem 3.3 are generalized to the case of sets $V_{r,a} = \{x \in H : w(x+a) < r\}$, where $w : H \to \mathbb{R}$ is a polynomial, second-degree and under appropriate conditions, the estimate

$$R = O_s\left((1 + ||a||^{3k+3})(n^{-s} + E||X_t||^2 + n^{-(k+1)/2}E||X^t||^{k+3})\right) \tag{3.9}$$

is obtained, valid for every $s > 0$. We recall that $X^t = X\chi\{||X|| \le n^{1/2}\}$ and $X_t = X - X^t$. Sazonov and Zalesskii (1985) developed truncation techniques adapted to the nonuniform estimates of the remainder. Using these techniques Bentkus and Zalesskii (1985) inserted the factor $(1+\rho)^{-m}$, $m > 0$, in the estimate (3.9) provided $E||X||^m < \infty$, where ρ denotes the distance between $0 \in H$ and the boundary of the set $V_{r,a}$. Nagaev and Chebotarev (1986, 1987, 1989a, b) considerably sharpened the estimate of the remainder in Theorems 3.1 and 3.3. These papers contain also the Bergström expansions in Hilbert space (see also Bentkus (1984e) and Bentkus and Zalesskii (1985) for expansions of an intermediate type). Papers of Sazonov, Ul'yanov and Zalesskii (1987a, b), Sazonov and Ul'yanov (1991) are devoted to a detailed investigation of uniform and nonuniform estimates of the remainders in Theorems 3.1 and 3.3. The case of non-identically distributed summands was investigated by Bentkus (1984e). Ul'yanov (1987) removed certain unnecessary restrictions from Bentkus (1984e) and obtained a more exact estimate of the remainder. Asymptotic expansions are also treated by Koroliuk and Borovskikh (1984).

We now state one result concerning asymptotic expansions in Banach spaces. Let $A_r = \{x \in B : F(x) < r\}$, where F is a functional $F : B \to \mathbb{R}$. Let

$$\sigma_{k+1}^2 = E_{k+1}(F^{(k+1)}(Y)Y_1 \cdots Y_{k+1})^2,$$

where the symbol E_{k+1} means that the expectation is taken only with respect to the r.e. Y_{k+1}.

Theorem 3.4. *Suppose that a r.e. $X \in B$ satisfies the CLT. Let a number $\gamma > 0$ and an integer $k \geq 0$ be fixed. Assume that F is $3k+3$ times Fréchet differentiable and that there exists a constant $M \geq 0$ such that*

$$\sup_{x \in B}(1 + ||x||)^{-M}||F^{(j)}(x)|| < \infty$$

for all $j = 0, \dots, 3k+3$. If

$$P\{\sigma_{k+1}^2 < \delta\} = O(\delta^m)$$

for sufficiently large $m = m(\varepsilon, k)$ as $\delta \downarrow 0$, then $E||X||^{k+2} < \infty$ ensures the existence of the asymptotic expansion

$$P\{S_n \in A_r\} = P\{Y \in A_r\} + \sum_{s=1}^{k} \varepsilon^s EP_s\nu(A_r) + R.$$

Furthermore, if $E||X||^{k+3} < \infty$, then $R = O(\varepsilon^{k+1-\gamma})$.

The theorem is contained in Götze (1989), where more general and more exact results are obtained. Earlier Götze (1983) proved the theorem for $k = 0$. In Götze (1984) the theorem is proved for $F(x) = ||x||^{k+1}$ and $B = L_{k+1}$. Vinogradova (1985) showed that the theorem is valid for symmetric F and $k = 1$.

§2.4. Asymptotic Expansions in the Local Limit Theorem

Let us consider the following Cramér-type condition: there exist a non-negative operator $K : H \to H$ and constants $\rho < \infty$, $\delta > 0$ such that for $r > 0$

$$\sup\{|E\exp(i(x, X))| : (Kx, x) \geq r^2\} \leq \rho r^{-\delta}. \tag{4.1}$$

Clearly, this is stronger than the Cramér condition (3.6) of the previous section. Put $A_{a,r} = \{x \in H : ||x + a||^2 < r\}$.

Theorem 4.1. *Suppose that $E||X||^{k+2} < \infty$ for some $k = 0, 1, \dots$ and that for some $\varepsilon > 0$ the integral*

$$J_\varepsilon = \int_{|t| \geq n^{1-\varepsilon}} |E\exp(it||S_n + a||^2)|dt$$

exists. Then the asymptotic expansion

$$\frac{d}{dr}P\{S_n \in A_{a,r}\} = \frac{d}{dr}P\{Y \in A_{a,r}\} + \sum_{s=1}^{k} \varepsilon^s \frac{d}{dr}EP_s(A_{a,r}) + R$$

is well defined. If

$$J_\varepsilon = O((1 + ||a||^{k+3})n^{-(k+1)/2}) \tag{4.2}$$

and $E||X||^{k+3} < \infty$, then

$$R = O((1 + ||a||^{k+3})n^{-(k+1)/2}). \tag{4.3}$$

If Cramér's condition (4.1) holds with an operator $K \geq 0$ such that the range of the operator K cov X is of sufficiently large or infinite dimension, then the integral J_ε exists for n sufficiently large, (4.2) holds and the remainder R therefore admits the estimate (4.3).

From this result one can also derive an asymptotic expansion for $P\{||S_n + a|| < r\}$. The theorem is a consequence of more general and more exact results due to Bentkus (1985a). Earlier Chebotarev (1982) proved under certain conditions that

$$\left|\frac{d}{dr}\{P\{||S_n||^2 < r\} - P\{||Y||^2 < r\}\}\right| = O(n^{-1/2})$$

if $E||X||^3 < \infty$ and the coordinates of X are independent. Recently Bentkus and Zitikis (to appear elsewhere) have generalized Theorem 4.1 to the case of a Banach space B and sets $A_{a,r} = \{x \in B : w(a + x) < r\}$, where $w : B \to \mathbb{R}$ is a polynomial of arbitrary degree.

Chapter 3
Applications

§3.1. Cramér–von Mises Statistics

This section considers the statistics

$$\omega_n^p(q) = n^{p/2} \int_0^1 |F_n(t) - t|^p q(t) dt. \tag{1.1}$$

Here $p > 0$ is an integer, $q : [0,1] \to [0,\infty)$ denotes a Lebesgue measurable weight function, and $F_n(t)$ denotes the empirical distribution function based on an independent random sample $x_1, \dots, x_n$ from the uniform distribution on $[0,1]$,

$$F_n(t) = n^{-1} \sum_{i=1}^{n} \mathcal{X}\{x_i < t\}.$$

We shall use the name "the Cramér–von Mises" statistics for $\omega_n^p(q)$. Denote the corresponding distribution functions by

$$U_n^p(x;q) = P\{\omega_n^p(q) < x\},$$

and let

$$U^p(x;q) = \lim_{n\to\infty} U_n^p(x;q).$$

Note that the statistics of the type

$$n^{p/2} \int_{-\infty}^{\infty} |F_n(t) - F(t)|^p q(F(t)) dF(t),$$

based on an arbitrary continuous distribution function F, can be reduced to (1.1) by changing variables.

We start with the discussion of results concerning the well-known ω_n^2-test,

$$\omega_n^2 = n \int_0^1 (F_n(t) - t)^2 dt \tag{1.2}$$

since in this case we are able to derive an almost complete set of convergence rate results from the general results obtained in Banach spaces. It is well known (see (Smirnov (1937)) that ω_n^2 converge weakly to

$$\omega^2 = \int_0^1 (W(t) - tW(1))^2 dt,$$

where $W(t)$, $0 \le t \le 1$, is the standard Wiener process. Prokhorov and Sazonov (1969) had noted that the following representation holds:

$$\omega_n^2 = ||S_n||^2, \quad S_n = n^{-1/2}(X_1 + \cdots + X_n). \tag{1.3}$$

Here $X, X_1, \ldots \in L_2(0,1)$ is a sequence of i.i.d. r.v.'s taking values in the Hilbert space $L_2(0,1)$ and

$$X(t) = \mathcal{X}\{x < t\} - t, \tag{1.4}$$

where x is a r.v. uniformly distributed in $[0,1]$. Clearly the X in (1.4) as a r.e. in $L_2(0,1)$ has mean zero and is bounded: $||X|| \le 1/3$. Therefore we may apply the general results obtained in Hilbert space. Indeed, the following theorem is a consequence of these general results (it is sufficient to apply the result due to Götze (1979)).

Theorem 1.1. *There exists an absolute constant C such that*

$$\Delta_n := \sup_{x\in\mathbb{R}} |U_n^2(x) - U^2(x)| \le Cn^{-1}, \tag{1.5}$$

where $U_n^2(x) = P\{\omega_n^2 < x\}$ and $U^2(x) = P\{\omega^2 < x\}$. Moreover,

$$U_n^2(x) = U^2(x) + \sum_{s=1}^{k} a_s(x) n^{-s} + R_k, \tag{1.6}$$

where the $a_s(x)$ are certain known functions and $|R_k| \le C(k)n^{-k-1}$.

The proof of Theorem 1.1 depends strongly on the following estimate. For sufficiently large positive A,

$$|E \exp(it\omega_n^2)| \le C(k,A)(1+|t|)^{-A} \tag{1.7}$$

for $|t| < n$ and $k = 0$ in case (1.5), and for $|t| \le n^{k+1}$ in case (1.6). Now it is known (see Bentkus and Zitikis (1988)) that for every positive A there exist constants $C_1(A)$, $C_2(A)$ such that

$$|E \exp(it\omega_n^2)| \le C_1(A)(1+|t|)^{-A} \tag{1.8}$$

for all $t \in \mathbb{R}$ and $n \ge C_2(A)$.

The remainder term R_k in (1.6) is nonuniform in $|x|$, i.e. for all $A > 0$

$$|R_k| = |R_k(x)| \le C(k,A)n^{-k-1}(1+|x|)^{-A}.$$

This follows from known estimates similar to (1.7) for the derivatives of the characteristic function $E\exp(it\omega_n^2)$ (see, e.g., Bentkus and Zitikis (1988)) and general results in Hilbert space concerning nonuniform estimates (see Sazonov and Zalesskii (1985), Bentkus and Zalesskii (1985)).

It should be noted that the proof of Theorem 1.1 has a long history. Papers by Sazonov (1968, 1969), Rosenkratz (1969), Kiefer (1972), Nikitin (1972), Orlov (1974) and Csörgő (1976) contain results of the type $\Delta_n = O(n^{-\beta})$ with various $\beta < 1/2$. The statistics ω_n^2 were also investigated by Koroliuk and Borovskikh (1984).

Theorem 1.2. *There exist absolute positive constants* C_1, C_2 *and a function* $I(n,t)$ (an analog of the classical "Cramér series" in large deviations results) *such that*

$$1 - U_n^2(x) = (1 - U^2(x))I(n,x)(1 + C_1\theta(1+\sqrt{x})/\sqrt{n}) \qquad (1.9)$$

when $0 \le x \le C_s n$. *Furthermore, if* $x \le C_2 n^{1/3}$ *then*

$$1 - U_n^2(x) = (1 - U^2(x))(1 + C_1\theta(1 + x^{3/2})/\sqrt{n}). \qquad (1.10)$$

The quantity θ *satisfies* $|\theta| \le 1$.

The pioneer papers of Osipov (1977, 1978a, 1978b) had a strong influence on the investigations of large deviations for ω_n^2 and in Hilbert space, as well. Theorem 1.2 is a special case of a general result in Hilbert space due to Yurinskii (1988). That paper contains a construction of the Cramér series $I(n,t)$. Inequality (1.10) also follows from Zalesskii (1989), Račkauskas (1988) and with a less precise estimate $O((1+\sqrt{x})n^{-1/6})$ of the remainder term from Bentkus (1986c).

Recall that C_b^k denotes the class of functions having k continuous and bounded derivatives.

Theorem 1.3 (Bentkus and Zitikis (1988)). *Let* $k = n/2 - 1$ *if* n *is even and* $k = (n-1)/2$ *if odd. The distribution function* $U_n^2 \in C_b^k$, *but* $U_n^2 \notin C_b^{k+1}$. *Moreover, for all* $m > 0$, $p = 0, 1, \ldots, n \ge 2(p+1)$ *and* $k = 0, 1, \ldots$,

$$\sup_{x\ge 0}(1+x^m)\left|\left(\frac{d}{dx}\right)^p\left\{U_n^2(x) - U^2(x) - \sum_{s=1}^k a_s(x)n^{-s}\right\}\right| \le C(m,p,k)n^{-k-1}.$$

The proof of Theorem 1.3 (i.e., the local limit theorem for ω_n^2 with an asymptotic expansion) is based on a detailed analysis of the characteristic functions and strongly depends on estimates such as (1.8). The result that $U_n^2 \in C_b^k$, $U_n^2 \notin C_b^{k+1}$, slightly improves the corresponding result due to Csörgő and Stacho (1980).

The condition $E\|X\|^2 < \infty$ is necessary and sufficient for the CLT in Hilbert space (see, e.g., Araujo and Giné (1980)). In the case of $\omega_n^2(q)$ the condition

$$\int_0^1 t(1-t)q(t)dt < \infty \qquad (1.11)$$

is clearly equivalent to $E\|X\|^2 < \infty$ for the r.e. X in (1.4) considered as an element of the space

$$L_2(q) = L_2((0,1), q(t)dt), \ \|x\|^2 = \int_0^1 x^2(t)q(t)dt.$$

Thus (1.11) guarantees the existence of

$$\lim_{n\to\infty} U_n^2(x,q) = U^2(x,q).$$

This result is wellknown and was proved by Chibisov (1964) using other methods. If (1.11) is fulfilled, then the condition $E\|X\|^{2m} < \infty$, $m \ge 0$, is equivalent to (see, e.g., Zitikis (1989))

$$\int_0^1 \left\{\int_0^x tq(t)dt\right\}^m dx < \infty,$$

$$\int_0^1 \left\{\int_x^1 (1-t)q(t)dt\right\}^m dx < \infty. \qquad (1.12)$$

On the other hand, one can easily show that the existence of the integrals

$$\int_0^{1/2} \left(\int_x^{1/2} q(t)dt\right)^m dx < \infty,$$

$$\int_{1/2}^1 \left(\int_{1/2}^x q(t)dt\right)^m dx < \infty \qquad (1.13)$$

is equivalent to (1.11) and (1.12) if $m \ge 1$. For example, if the function q is symmetric about the point $t = 1/2$ and $q(t) = t^{-\delta}$ with some δ for $0 < t \le 1/2$ then (1.13) is fulfilled if and only if $\delta < 1 + 1/m$.

Theorem 1.4. *Suppose that* (1.13) *holds for some* $1 \le m < 2$. *Then*

$$\Delta_n(q) := \sup_{x\ge 0}|U_n^2(x;q) - U^2(x;q)| \le C(q,m)n^{1-m}.$$

Moreover, $\Delta_n(q) = o(n^{1-m})$ *as* $n \to \infty$.

Theorem 1.4 follows from the known estimates of the convergence rate in Hilbert space (see Götze (1979), Zalesskii (1982)). One has to verify that the corresponding r.e. $X \in L_2(q)$, or equivalently, the limiting Gaussian process $Y(t) = W(t) - tW(1)$, where $W(t)$, $t \in [0,1]$, is the standard Wiener process, is not concentrated in a finite-dimensional subspace of $L_2(q)$. It suffices to show that W is not finite-dimensional since $tW(1)$ is concentrated in a one-dimensional subspace. Clearly, to this end it is enough to construct linear measurable functionals $l_1, \ldots, l_m : L_2(q) \to \mathbb{R}$ for every integer $m \ge 1$ such that the Gaussian r.v.'s $l_1(W), \ldots, l_m(W)$ are independent and nondegenerate. Without loss of generality we may assume that $meas\{t : q(t) > 0\} > 0$. Therefore there exist pairwise disjoint intervals $(a_i, b_i) \subset (0,1)$, $1 \le i \le m$, such that $\int_{a_i}^{b_i} q(t)dt > 0$. Thus we can put

$$l_i(W) = \int_{a_i}^{b_i} (W(t) - W(a_i))q(t)dt.$$

Consider the condition

$$\inf\{q(t) : t \in (0,1)\} > 0. \qquad (1.14)$$

We shall assume that q satisfies the following condition: there exist a finite number of points $t_0 = 0 < t_1 < \cdots < t_N < t_{N+1} = 1$ such that the function q is monotone and either convex or concave on every interval (t_{i-1}, t_i). Moreover, we suppose that there exist numbers $\alpha \geq 0$ and $c(\alpha) < \infty$ such that

$$\max\{q(t_{i-1}+\varepsilon);\ q(t_i-\varepsilon)\} \leq c(\alpha)\varepsilon^{-\alpha} \tag{1.15}$$

for all $\varepsilon > 0$ such that $2\varepsilon < t_i - t_{i-1}$.

Theorem 1.5 (Zitikis (1989)). *Suppose that conditions* (1.13) *with* $m=2$ *and* (1.14), (1.15) *hold. Then*

$$\Delta_n(q) \leq c(q)n^{-1}.$$

Moreover, if (1.13) *holds with* $m = \mu(1+\varepsilon) \geq 2$ *for some* $\varepsilon > 0$ *and* $\mu = 1, 2, \ldots$, *then*

$$\sup_{x>0} x^{\mu}|U_n^2(x;q) - U^2(x;q)| \leq c(q,\varepsilon,\mu)n^{-1}.$$

The paper of Zitikis (1989) also contains asymptotic expansions for $U_n^2(\cdot;q)$ and its derivatives. Estimates of the remainder term obtained are uniform and nonuniform. These expansions are found under the additional condition that q' satisfies (1.15).

Concerning the large deviations results, one can replace U_n^2 by $U_n^2(\cdot;q)$ in Theorem 1.2 if the following condition is satisfied: there exists a positive C such that

$$\int_{1/2}^{1} \exp\left(C\left[\int_{1/2}^{\tau} q(t)dt\right]^{1/2}\right) d\tau < \infty,$$

$$\int_{0}^{1/2} \exp\left(C\left[\int_{\tau}^{1/2} q(t)dt\right]^{1/2}\right) d\tau < \infty. \tag{1.16}$$

This replacement is possible since the general result of Yurinskii (1988) is valid when $E\exp(c||X||) < \infty$ for some $c > 0$, which in our case is equivalent to (1.16).

Let us now discuss the situation when $p \neq 2$. Similarly to the representation (1.3), one can write

$$\omega_n^p(q) = ||S_n||^p \tag{1.17}$$

with the same r.e. X of (1.4) but considered as an element of the space $L_p = L_p(q)$, $p \geq 1$, of functions $f : [0,1] \to \mathbb{R}$ with finite integral

$$||f||^p = \int_0^1 |f(t)|^p q(t)dt.$$

The weak convergence of the distributions $U_n^p(x;q)$ to $U^p(x;q)$ was investigated by Csörgő and Horvath (1988) and Norvaiša (1990). Corresponding deep and general results on the convergence rate in the CLT in Banach spaces (see, e.g., Götze (1989), (1984) and Zalesskii (1988)) were obtained under certain "variance" conditions (see § 1.3). Unfortunately up till now this condition still has not been expressed in terms of the weight function q, with the exception of certain special cases, for example, $q(t) \equiv 1$. In this particular case we have the following result.

Theorem 1.6 (Götze (1984)). *Suppose that* $p \geq 4$ *is even and that* $q(t) \equiv 1$. *Then*

$$\sup_{x\geq 0} |U_n^p(x;1) - U^p(x;1)| \leq C(p)n^{-1}.$$

Moreover, one can expand $U_n^p(x;1)$ *in an asymptotic series with a remainder term of the order* $O(n^{\varepsilon-p/2})$, $\varepsilon > 0$.

For further details and results, see Götze (1984). Thus, for $U_n^p(x,q)$, $p \neq 2$, we have the Berry–Esseen estimates and large deviations results with a remainder term of order $O(n^{-1/6})$ only.

Theorem 1.7. *If* $p \geq 3$ *and*

$$\int_0^1 t^p(1-t)^p q(t)dt < \infty,$$

$$\int_{1/2}^1 \left(\int_{1/2}^x q(t)dt\right)^{3/p} dx < \infty,$$

$$\int_0^{1/2} \left(\int_x^{1/2} q(t)dt\right)^{3/p} dx < \infty, \tag{1.18}$$

then

$$\sup_{x\geq 0}(1+x^{3/p})|U_n^p(x;q) - U^p(x;q)| \leq C(q)n^{-1/6}. \tag{1.19}$$

When $1 < p < 3$, *estimate* (1.19) *is valid if there exists* $\alpha > 0$, $\alpha < p/3$, *such that*

$$\int_0^1 t^{\alpha}(1-t)^{\alpha} q(t)dt < \infty.$$

The theorem is found in Paulauskas and Račkauskas (1991) for $p \in (1,3)$. For $p \geq 3$ the result follows from Paulauskas and Račkauskas (1991) provided that the distribution function $x \to (U^p(x;q))^{1/p}$ has a bounded density and that the corresponding random variable $X \in L_p$ has a third moment. The first condition follows from general results of Davydov and Lifshits (1984) and Rhee and Talagrand (1986). The second one is a consequence of the identity

$$E\phi(||X||) = \int_0^1 \phi\left(\left\{\int_0^x t^p q(t)dt + \int_x^1 (1-t)^p q(t)dt\right\}^{1/p}\right) dx, \tag{1.20}$$

valid for every $\phi : [0,\infty) \to [0,\infty)$. When the function $\phi(t) = t^3$, (1.20) is equivalent to (1.18).

The following theorem is a consequence of more general results due to Bentkus and Račkauskas (1990) (if $1 < p < 3$, one must use additionally the result of Paulauskas and Račkauskas (1991) about differentiable functions in L_p).

Theorem 1.8. *If*

$$\int_0^1 t(1-t)q(t)dt < \infty$$

when $p \geq 3$ *and*

$$\int_0^1 t^\alpha (1-t)^\alpha q(t)dt < \infty$$

for some $\alpha > 0$, $\alpha < p/3$, *when* $1 < p < 3$, *then the limiting distribution function* $U^p(x;q)$ *exists and the following three statements are equivalent:*

(a) *there exists* $h > 0$ *such that*

$$\int_{1/2}^1 \exp\left(h\left[\int_{1/2}^x q(t)dt\right]^{1/2p}\right)dx < \infty,$$

$$\int_0^{1/2} \exp\left(h\left[\int_x^{1/2} q(t)dt\right]^{1/2p}\right)dx < \infty;$$

(b) *there exist constants* $M_i = M_i(p,q) > 0$, $i = 1,2$, *such that*

$$\left|\frac{1-U_n^p(x;q)}{1-U^p(x;q)} - 1\right| \leq M_1(1+x^{1/p})n^{-1/6},$$

when $0 \leq x \leq M_2 n^{p/6}$;

(c) *for each function* $f : \mathbb{R} \to \mathbb{R}$ *such that* $f(n) \to 0$, *as* $n \to \infty$,

$$\frac{1-U_n^p(x;q)}{1-U^p(x;q)} \to 1 \quad \text{as } n \to \infty,$$

uniformly for $0 \leq x \leq f(n)n^{p/6}$.

§3.2. *L*-Statistics

This section discusses only recent work on large deviations for *L*-statistics since these results were obtained by reducing the problem to results in Banach spaces or by applying appropriate techniques. We refer to Stigler (1974), Serfling (1980), Helmers (1982), Bhattacharya and Denker (1990), Norvaiša and Zitikis (1991), and Zitikis (1991a) for other results concerning *L*-statistics.

Let $X_1,\dots,X_n$ be independent copies of a r.v. $X \in \mathbb{R}$ with distribution function F. Consider the *L*-statistic

$$\ell_n = n^{-1}\sum_{i=1}^n c_{in}X_{i:n},$$

where $X_{1:n} \leq \cdots \leq X_{n:n}$ are the order statistics of $X_1,\dots,X_n$ and $c_{1n},\dots,c_{nn}$ are certain real coefficients.

When the coefficients are generated by a weight function $J : [0,1] \to \mathbb{R}$ (see, e.g., Stigler (1974)), that is

$$c_{in} = c_{in}^0 = n\int_{(i-1)/n}^{i/n} J(u)du,$$

we shall denote the corresponding *L*-statistic by ℓ_n^0. We shall assume that there exists an $L \geq 0$ such that

$$|J(u) - J(v)| \leq L|u-v|$$

for all $u, v \in [0,1]$. It is known (see Stigler (1974)) that if $EX^2 < \infty$, then there exist finite limits

$$\mu = \lim_{n\to\infty} E\ell_n^0, \quad \sigma^2 = \lim_{n\to\infty} nE(\ell_n^0 - \mu)^2,$$

where

$$\mu = \int_{-\infty}^{\infty} xJ(F(x))dF(x),$$

$$\sigma^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(F(x))J(F(y))\{F(x\wedge y) - F(x)F(y)\}dxdy.$$

We shall assume that $\sigma > 0$. This guarantees that the limiting distribution of the r.v. $n^{-1/2}(\ell_n^0 - \mu)/\sigma$ is the standard normal distribution $\Phi(x)$.

Theorem 2.1. *If for some* $h > 0$

$$E\exp(h|X|^{1/2}) < \infty, \tag{2.1}$$

then for each sequence $b_n \to 0$

$$P\{\sqrt{n}(\ell_n^0 - \mu) > \sigma x\}/[1-\Phi(x)] \to 1 \quad \text{as } n \to \infty, \tag{2.2}$$

uniformly for $0 \leq x \leq b_n n^{1/6}$.

Note that the theorem (and the results below) remains valid for the quotient

$$P\{\sqrt{n}(\ell_n^0 - \mu) < \sigma x\}/\Phi(x)$$

when $-b_n n^{1/6} \leq x \leq 0$. Indeed, it is sufficient to replace J by $-J$.

Assertions (2.1) and (2.2) are equivalent when $J \equiv const. \neq 0$. In this case the statistic ℓ_n^0 reduces to a sum of i.i.d. r.v.'s and the equivalence is proved, for example, in Bentkus and Račkauskas (1990). The same remark applies to Theorem 2.2 below.

Let us fix a number a, $0 < a \leq 1/2$, and put $x_n = n^{a/(4-2a)}$.

Theorem 2.2. *If there exists* $h > 0$ *such that*

$$E\exp(h|X|^a) < \infty, \tag{2.3}$$

then for each sequence $b_n \to 0$

$$P\{\sqrt{n}(\ell_n^0 - \mu) > \sigma x\}/[1 - \Phi(x)] \to 1 \quad \text{as } n \to \infty,$$

uniformly for $0 \le x \le b_n x_n$.

Theorem 2.1 and Theorem 2.2 are stable with respect to small perturbations of the coefficients $c_{11}^0, \ldots, c_{nn}^0$ in the following sense.

Theorem 2.3. *Theorems* 2.1 *and* 2.2 *remain valid if the coefficients* $c_{1n}^0, \ldots, c_{nn}^0$ *are replaced by* $c_{1n}, \ldots, c_{nn}$ *such that*

$$\sum_{i=1}^n |c_{in} - c_{in}^0|^3 = O(n^{-1}) \tag{2.4}$$

in the case of Theorem 2.1 *and*

$$\sum_{i=1}^n |c_{in} - c_{in}^0|^{(2-a)/(1-a)} = O(n^{a/(a-1)}) \tag{2.5}$$

in the case of Theorem 2.2. *In both cases the statistic* ℓ_n^0 *should be replaced by* ℓ_n.

The application of Hölder's inequality shows that (2.4) implies (2.5). Theorems 2.1–2.3 were proved by Bentkus and Zitikis (1990). The proofs are based on the representation $\ell_n = S_n + \omega_n^2 + R$, where S_n denotes a sum of i.i.d. $\mathbb{R}$-valued r.v.'s, ω_n^2 is an ω^2-statistic and R is a remainder term. The estimation of S_n is based on the well-known one-dimensional results (see, e.g., Ibragimov and Linnik (1965), Petrov (1972), and Bentkus and Račkauskas (1990)). The estimation of ω_n^2 is based on large deviations results in Hilbert space by Bentkus and Račkauskas (1990) (see also § 3.1). Using the same techniques Zitikis (1991b) showed that the speed of convergence in Theorem 2.1 and 2.2 is of the order $(1 + x^3)n^{-1/2}\ln n$. Combining the methods and techniques developed to obtain convergence rates and large deviations theorems in Banach spaces, Zitikis (1991a) proved the next result.

Theorem 2.4. *Condition* (2.3) *guarantees that there exist constants* $C = C(a, J)$, *and* $A = A(a, J) > 0$ *such that*

$$|1 - P\{\sqrt{n}(\ell_n^0 - \mu) > \sigma x\}/[1 - \Phi(x)]| \le C(1 + x^3)n^{-1/2}$$

for $0 \le x \le A x_n$.

Cramér-type large deviations for L-statistics were investigated by Vandamaele and Veraverbeke (1982), Puri and Seoh (1987), and Aleškevičienė (1989).

§3.3. Kolmogorov–Smirnov Statistics

Let F_n be the empirical distribution function based on a sample $x_1, \ldots, x_n$ of observations from a uniform distribution on $[0,1]$ and let

$$D_n(t) = n^{1/2}(F_n(t) - t).$$

In the previous sections we have seen that it is useful to consider $\{D_n, n \ge 1\}$ as a sequence of r.e.'s in L_p. Here we shall consider D_n as a r.e. in the Skorokhod space $D[0,1]$,

$$D_n(t) = n^{-1/2}\sum_{i=1}^n X_i(t),$$

where X, X_1, $X_2, \ldots$ are i.i.d. r.e.'s in $D[0,1]$ having the same distribution as the process $X(t) = \mathcal{X}\{x < t\} - t$, $t \in [0,1]$, with r.v. x uniformly distributed on $[0,1]$. Therefore the weak convergence of D_n (that is, the classical Donsker theorem) is nothing more than the CLT in $D[0,1]$ for i.i.d. summands having the special structure. Any of Theorems 4.7–4.10 of Chapter 1 can be easily applied to D_n. For example, to apply Theorem 4.7 the following two conditions are sufficient:

$$E(X(t) - X(s))^2 \le t - s,$$

$$E(X(u) - X(t))^2(X(t) - X(s))^2 \le (u - s)^3, \ 0 \le s \le t \le u \le 1.$$

Of course, this application has only methodological importance. Here we shall apply the results of § 1.4 to weighted empirical processes. To this end some new definitions are required. Let Q be the class of functions $q : [0,1] \to \mathbb{R}^+$ continuous and increasing on $[0,1/2]$ and symmetric about the point $1/2$: $q(t) = q(1-t)$, $0 \le t \le 1/2$. Let

$$W_0(t) = W(t) - tW(1), \ 0 \le t \le 1,$$

be the Brownian bridge, where W is the standard Wiener process. The weighted empirical process and the weighted Brownian bridge are given by

$$D_{n,q}(t) = D_n(t)/q(t) = n^{-1/2}\sum_{i=1}^n X_{i,q}(t),$$

$$W_{0,q}(t) = W_0(t)/q(t),$$

respectively, where $X_{i,q}(t) = X_i(t)/q(t)$. Chibisov (1964) (see also O'Reilly (1974)) proved that for $q \in Q$ the distributions $\mathcal{L}(D_{n,q})$ converge weakly to $\mathcal{L}(W_{0,q})$ if and only if

$$\int_0^{1/2} \exp(-\varepsilon q^2(t)/t)t^{-1}dt < \infty \tag{3.1}$$

for all $\varepsilon > 0$. We shall compare this result with the one obtained by applying Theorem 4.8 of Chapter 1. Suppose in addition that $t^{1/2}/q(t)$ is a nondecreasing function. Performing some simple estimates (for details, see Paulauskas and Stieve (1990)), we have for all $0 \le s \le t \le u \le 1$

$$E(X_q(t) - X_q(s))^2 \le C \int_s^t v^{-1/2}/q(v)dv, \tag{3.2}$$

$$E\left((X_q(t) - X_q(s))^2 \wedge 1\right)(X_q(u) - X_q(t))^2 \le C\left(\int_s^u v^{-1/2}/q(v)dv\right)^2 \tag{3.3}$$

where $X_q(t) = X(t)/q(t)$. Therefore, if

$$\int_0^{1/2} v^{-1/2}/q(v)dv < \infty, \tag{3.4}$$

then the conditions of Theorem 4.8 are satisfied with $\alpha_1 = 1$, $\alpha_2 = 2$ and $F_1(t) = F_2(t) = C\int_0^t v^{-1/2}/q(v)dv$ and this implies the weak convergence of $\mathcal{L}(D_{n,q})$. It follows that condition (3.4) is very close to the optimal O'Reilly–Chibisov condition.

We close this section with a result implied by Theorem 4.11 of Chapter 1 applied to weighted empirical processes. Let

$$\Delta_{n,q}(x) = |P\{||D_{n,q}||_\infty < x\} - P\{||W_{0,q}||_\infty < x\}|.$$

Theorem 3.1 (Paulauskas and Stieve (1990)). *Suppose that $q \in Q$ and for some $0 \le \delta \le 2/3$ and $0 < \gamma < 1$ the function $t^{1/(3+\delta)}/q(t)$ is increasing and*

$$\sup_{0 \le t \le 1/2} t^{(1-\gamma)/(3+\delta)}/q(t) < C.$$

Then there exists a constant $C_0 = C_0(q)$ such that

$$\Delta_{n,q}(x) \le C_0 n^{-(1+\delta)/10}(1+x)^{-3}(\ln n + \ln(1+x))^2. \tag{3.5}$$

To prove this one needs to verify the conditions of Theorem 4.12 of Chapter 1, among them condition (4.3) for $W_{0,q}$. See Paulauskas (1990) and Paulauskas and Stieve (1990) for details.

Here it is necessary to note that there exists a vast literature on weighted empirical and quantile processes (see, for example, Csörgő et al. (1986) and the references therein). Most of these papers employ the so-called Hungarian (or KMT) construction. Using these results, one can easily deduce the following: if $q(t) = t^\alpha$, $1/4 < \alpha < 1/2$, and condition (4.3) is fulfilled for $W_{0,q}$, then

$$\sup_{x \in \mathbb{R}} \Delta_{n,q}(x) \le Cn^{-1/2+\alpha}. \tag{3.6}$$

Hence the rate $n^{-1/6}$ in (3.6) corresponds to the rather natural weight function $t^{1/3}$ while in (3.5) for the same rate we have to choose the function $t^{3/11}$.

§3.4. Empirical Processes

In this section we shall state two estimates of the convergence rate in the CLT for general empirical processes. There exist a large number of papers devoted to the CLT itself. We refer the reader to survey papers by Dudley (1978), Gaensler and Stute (1979), and Giné and Zinn (1984).

Let $X, X_1, X_2, \ldots$ be i.i.d. r.e.'s taking values in a measurable space $(\mathcal{X}, \mathcal{A})$. Let $\mu = \mathcal{L}(X)$ and $\mu f = \mu(f) = \int_{\mathcal{X}} f(x)\mu(dx)$. Let μ_n and E_n denote the empirical measure and empirical process, respectively, associated with μ, i.e.,

$$\mu_n = n^{-1}\sum_{i=1}^n \delta_{X_i}, \quad E_n = \sqrt{n}(\mu_n - \mu),$$

where δ_x denotes the point mass at $x \in \mathcal{X}$. We shall consider the empirical process $E_n(f)$, $f \in \mathcal{F}$, indexed by some class $\mathcal{F} \subset L_2(\mathcal{X}, \mathcal{A}, \mu)$ of measurable functions $f : \mathcal{X} \to \mathbb{R}$. Let us define the pseudo-metrics ρ_μ and $e_{\mu,p}$, $1 \le p \le \infty$, on $\mathcal{F}$ as follows:

$$\rho_\mu(f,g) = (\mu(f-g)^2 - \mu^2(f-g))^{1/2},$$

$$e_{\mu,p}(f,g) = (\mu|f-g|^p)^{1/p}, \text{ if } 1 \le p < \infty,$$

$$e_{\mu,\infty}(f,g) = ess\ sup|f-g|,$$

where the *ess sup* is taken with respect to the measure μ. We suppose that the class $\mathcal{F}$ is μ-pre-Gaussian. This means that there exists a zero-mean Gaussian process $B_\mu(f)$, $f \in \mathcal{F}$, with covariance

$$EB_\mu(f)B_\mu(g) = \mu(fg) - \mu(f)\mu(g), \quad f, g \in \mathcal{F},$$

which has a version with bounded and uniformly continuous sample functions with respect to ρ_μ. In order to avoid measurability problems we suppose that the class $\mathcal{F}$ is countably determined (for μ). That is, we suppose that there exists a countable subclass $\mathcal{F}_0 \subset \mathcal{F}$ such that $||E_n||_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |E_n(f)| = ||E_n||_{\mathcal{F}_0}$ a.e. for all $n \ge 1$. If the class $\mathcal{F}$ is totally bounded under a pseudo-metric d, then the covering number $N_d(u)$ and the metric entropy $H_d(u)$ of $\mathcal{F}$ are defined as

$$N_d(u) = N(\mathcal{F}, d, u) := \min\{k : \exists f_1, \ldots, f_k \in \mathcal{F} : \min_{1 \le i \le k} d(f, f_i) < u \forall f \in \mathcal{F}\},$$

$$H_d(u) = \log N_d(u), \ u \ge 0.$$

The last notion that will be needed to formulate results is the Vapnik–Chervonenkis class. A class $\aleph$ of subsets of $\mathcal{X}$ is said to shatter a finite subset Γ_0 of $\mathcal{X}$ if every $\Gamma \subset \Gamma_0$ is of the form $C \cap \Gamma_0$ for some $C \in \aleph$. $\aleph$ is called a Vapnik–Chervonenkis class (VC for short) if for some $n \ge 1$ no n-element subset of $\mathcal{X}$ is shattered by $\aleph$. The least such n is called the index of $\aleph$ and is denoted by $v(\aleph)$. For a non-negative function $f : \mathcal{X} \to \mathbb{R}$, the set defined by $sub(f) := \{(x,t) \in \mathcal{X} \times \mathbb{R} : 0 \le 1 \le f(x)\}$ is called a subgraph of f. Let

$$D(\mathcal{F}) = \{sub(f) : f \in \mathcal{F}\}.$$

Following Dudley (1987), we call $\mathcal{F}$ a VC-subgraph class (for D) if $D(\mathcal{F})$ is a VC class.

As usual in order to estimate the convergence rate in the CLT, we need the ε-strip condition on the limiting Gaussian process: there exists a constant C such that for all $\varepsilon > 0$ and $r \ge 0$,

$$P\{r \le \|B_\mu\|_{\mathcal{F}} \le r + \varepsilon\} \le C\varepsilon(1+r)^{-3}. \tag{4.1}$$

Theorem 4.1 (Norvaiša and Paulauskas (1991)). *Let $\mathcal{F}$ be a countably determined VC subgraph class of functions f such that* $0 \le f \le 1$. *Assume that* (4.1) *holds. Then*

$$\begin{aligned} \Delta_n(\mathcal{F}, r) &:= |P\{\|E_n\|_{\mathcal{F}} < r\} - P\{\|B_\mu\| < r\}| \\ &= O((1+r)^{-3} n^{-1/6} \ln^2 n). \end{aligned} \tag{4.2}$$

Theorem 4.2 (Norvaiša and Paulauskas (1991)). *Let $\mathcal{F}$ be a countably determined class of functions f such that $0 \le f \le 1$. Assume that* (4.1) *is satisfied and that for some C and some $s \in (0, 1/2)$*

$$H(\mathcal{F}, e_{\mu,\infty}, u) \le C u^{-s}, \quad u > 0.$$

Then

$$\Delta_n(\mathcal{F}, r) = O((1+r)^{-3} n^{-(1-2s)/6} \ln^s n). \tag{4.3}$$

Let us consider as an example the multivariate empirical process. Namely let

$$\mathcal{X} = \mathbb{R}^d, \ (-\infty, t] = \{s = (s_1, \ldots, s_d) \in \mathbb{R}^d : s_1 \le t_1, \ldots, s_d \le t_d\},$$

$$\mathcal{F} = \{\chi_{(-\infty,t]}, \ t \in \mathbb{R}^d\}, \ E_n(t) = \sqrt{n}(F_n(t) - F(t)), \ t \in \mathbb{R}^d,$$

where $F(t)$ denotes a distribution function with law μ and F_n denotes the empirical distribution function related to F. Norvaiša and Paulauskas (1991) proved the validity of (4.1) under very mild conditions on F. Combining this with the known fact that the class of lower-left orthants of $\mathbb{R}^d$ is a VC class and is always countably determined (see Wenocur and Dudley (1981)), by Theorem 4.1 we obtain the following result.

Theorem 4.3. *Suppose there is an increasing sequence $t_k \in \mathbb{R}^d$, $k \ge 1$ such that the distribution function F satisfies* $0 < F(t_1) < F(t_2) < \cdots < 1$. *Then*

$$|P\{\sup_{t \in \mathbb{R}^d} |E_n(t)| < r\} - P\{\sup_{t \in \mathbb{R}^d} |B_F(t)| < r\}| = O((1+r)^{-3} n^{-1/6} \ln^2 n). \tag{4.4}$$

One can compare Theorem 4.3 with results which follow from the weak (or strong) invariance principle (see, e.g., Massart (1986, 1989) and references therein). Here we note only that in (4.4) the order of the remainder term does not depend on d, whereas usually the convergence rates obtained by the Hungarian construction are of order $n^{-1/2d}$. Hence (4.4) is better for $d > 3$. For more details on the comparison of these results and for a more complete list of references on this topic we refer the reader to the above-mentioned papers of Massart (1986, 1989) and Norvaiša and Paulauskas (1991).

References*

Acosta, A de, and Giné, E. (1979): Convergence of moments and related functionals in the general central limit theorem in Banach spaces. Z. Wahrscheinlichkeitstheorie Verw. Geb. **48**, 213–231. Zbl. 395.60008

Aleškevičienė, A. (1989): On large deviations for linear combinations of order statistics. Liet. Mat. Rink. (Litov. Mat. Sb.) **29**, 212–222 (in Russian). Zbl. 682.60021

Aliev, F.A. (1987): A lower bound for the convergence rate in the central limit theorem in Hilbert space. Teor. Veroyatn. Primen. **31**, No. 4, 825–828. English transl.: Theory Probab. Appl. **31**, No. 4, 730–733. Zbl. 623.60013

Aliev, F.A. (1989): On the convergence rate in the central limit theorem in Hilbert space. Teor. Veroyatn. Primen. **34**, No. 2, 407–409. English transl.: Theory Probab. Appl. **34**, No. 2, 361–363. Zbl. 672.60006

Araujo, A., and Giné, E. (1980): The Central Limit Theorem for Real and Banach-Valued Random Variables. John Wiley, New York. Zbl. 457.60001

Asriev, A.V., and Rotar', V.J. (1985): On the convergence rate in the infinite-dimensional central limit theorem for probabilities of hitting parallelepipeds. Teor. Veroyatn. Primen. **30**, No. 4, 652–661. English transl.: Theory Probab. Appl. **30**, No. 4, 691–701. Zbl. 586.60005

Averbukh, V.I., Smolyanov, O.G., and Fomin, S.V. (1971): Generalized functions and differential equations in linear spaces. Tr. Mosk. Mat. O.-va **24**, 133–174. English transl.: Trans. Mosc. Math. Soc. **24**, 140–184. Zbl. 234.28005

Badrikian, A., and Chevet, S. (1974): Measures Cylindriques. Espaces de Wiener et Functions Aléatoires Gaussiennes. Lect. Notes Math 379, Springer-Verlag, Berlin. Zbl. 288.60009

Barsov, S.S. (1985): On the accuracy of the normal approximation to the distribution of a random sum of random vectors. Teor. Veroyatn. Primen. **30**, No. 2, 351–354. English transl.: Theory Probab. Appl. **30**, No. 2, 376–379. Zbl.569.60021

Barsov, S.S. (1987): Rates of convergence to the normal distribution and decrease of the tail of the summand distribution. Teor. Veroyatn. Primen. **32**, No. 2, 356–358. English transl.: Theory Probab. Appl. **32**, No. 2, 329–331. Zbl. 623.60028

* For the convenience of the reader, references to reviews in Zentralblatt für Mathematik (Zbl.), compiled by means of the MATH database, and Jahrbuch über die Fortschritte der Mathematik (Jbuch) have, as far as possible, been included in this bibliography.

Bass, R., and Pyke, R. (1985): The space $D(A)$ and weak convergence for set-indexed processes. Ann. Probab. **13**, 860–864. Zbl. 585.60007

Bass, R., and Pyke, R. (1987): A central limit theorem for $D(A)$-valued processes. Stoch. Processes Appl. **24**, 109–131. Zbl. 617.60020

Bentkus, V. (1982): Estimates of the rate of convergence in the central limit theorem in the space $C(S)$. Dokl. Akad. Nauk SSSR **266**, 526–529. English transl.: Soviet Math. Dokl. **26**, 349–352. Zbl. 521.60005

Bentkus, V. (1984a): Lower bounds for the sharpness of a normal approximation in Banach spaces. Liet. Mat. Rink. **24**, No. 1, 12–18. English transl.: Lith. Math. J. **24**, No. 1, 6–10. Zbl. 584.60034

Bentkus, V. (1984b): Asymptotic analysis of the remainder term in the central limit theorem in Hilbert space. Liet. Mat. Rink **24**, No. 1, 5–11. English transl.: Lith. Math. J. **24**, No. 1, 2–6. Zbl. 586.60022

Bentkus, V. (1983): Differentiable functions defined in the spaces c_0 and $\mathbb{R}^k$. Liet. Mat. Rink. **24**, No. 2, 26–36. English transl.: Lith. Math. J. **24**, No. 2, 146–154. Zbl. 527.46036

Bentkus, V. (1984d): Asymptotic expansions in the central limit theorem in Hilbert space. Liet. Mat. Rink. **24**, No. 3, 29–50. English transl.: Lith. Math. J. **24**, No. 3, 210–225. Zbl. 568.60019

Bentkus, V., (1984e): Asymptotic expansions for the distributions of sums of independent random elements of a Hilbert space. Liet. Mat. Rink. **24**, No. 4, 29–48. English transl.: Lith. Math. J. **24**, No. 4, 305–319. Zbl. 573.60007

Bentkus, V. (1984f): The Asymptotic Analysis of Sums of Independent Banach-Space-Valued Random Elements. Doctoral Dissertation, Vilnius (in Russian).

Bentkus, V. (1984g): Asymptotic analysis of moments in the central limit theorem in Banach spaces. Liet. Mat. Rink. **24**, No. 2, 49–64. English transl.: Lith. Math. J. **24**, No. 2, 113–125. Zbl. 558.60007

Bentkus, V. (1985a): Asymptotic expansions in the local limit theorem in a Hilbert space. Liet. Mat. Rink. **25**, No. 1, 9–22. English transl.: Lith. Math. J. **25**, No. 1, 1–10. Zbl. 567.60003

Bentkus, V. (1985b): Concentration functions of sums of independent random elements of a Banach space. Liet. Mat. Rink. **25**, No. 2, 32–39. English transl.: Lith. Math. J. **25**, No. 2, 32–39. Zbl. 577.60005

Bentkus, V. (1986a): Dependence of the Berry-Esseen estimate on the dimension. Liet. Mat. Rink. **26**, No. 2, 205–210. English transl.: Lith. Math. J. **26**, No. 2, 110–114. Zbl. 612.60022

Bentkus, V. (1986b): Lower bounds for the rate of convergence in the central limit theorem in Banach spaces. Liet Mat. Rink. **25**, No. 4, 10–25. English transl.: Lith. Math. J. **26**, 312–320. Zbl. 588.60010

Bentkus, V. (1986c): Large deviations in Banach space. Teor. Veroyatn. Primen. **31**, No. 4, 710–716. English transl.: Theor. Probab. Appl. **31**, No. 4, 627–632. Zbl. 623.60012

Bentkus, V. (1986d): Asymptotic expansions for moments in the central limit theorem in Banach spaces. Liet. Mat. Rink. **26**, No. 1, 10–26. English transl.: Lith. Math. J. **26**, No. 1, 6–18. Zbl. 649.60024

Bentkus, V. (1987): Lower estimates of the convergence rate in the central limit theorem in Banach spaces. Probab. Theory Math. Stat. (Vilnius 1985). English transl.: VNU Science Press, Utrecht, The Netherlands **1**, 171–187. Zbl. 657.60013

Bentkus, V. (1990): Smooth approximations of the norm and differentiable functions with bounded support in Banach space. Liet. Mat. Rink. **30**, No. 3, 489–499. English transl.: Lith. Math. J. **30**, No. 3, 223–230. Zbl. 725.46009

Bentkus, V., and Liubinskas, K. (1987): Rate of convergence in the invariance principle in Banach spaces. Liet. Mat. Rink. **27**, No. 3, 423–434. English transl.: Lith. Math. J. **27**, No. 3, 205–213. Zbl. 635.60007

Bentkus, V., and Račkauskas, A. (1981): Convergence rate in the central limit theorem in infinite-dimensional spaces. Liet. Mat. Rink. **21**, No. 4, 9–18. English transl.: Lith. Math. J. **21**, No. 4, 271–276. Zbl. 493.60009

Bentkus, V., and Račkauskas, A. (1982): Estimates of the convergence rate of sums of independent random variables in a Banach space, I, II. Liet. Mat. Rink. **22**, No. 3, 12–28; No. 4, 8–20. English transl: Lith. Math. J. **22**, No. 3, 223–234; No. 4, 344–353. I. Zbl. 522.60007, II. Zbl. 522.60007

Bentkus, V., and Račkauskas, A. (1984): Estimates of the distance between sums of independent random elements in Banach spaces. Teor. Veroyatn. Primen. **29**, No. 1, 49–64. English transl.: Theory Probab. Appl. **29**, No. 1, 50–65. Zbl. 534.60008

Bentkus, V., and Račkauskas, A. (1990): On probabilities of large deviations. Probab. Theory Relat. Fields **86**, 131–154. Zbl. 678.60005

Bentkus, V., and Zalesskii, B. (1985): Asymptotic expansions with nonuniform remainders in the central limit theorem in Hilbert space. Liet. Mat. Rink. **25**, No. 3, 3–16. English transl.: Lith. Math. J. **25**, No. 3, 199–208. Zbl. 585.60011

Bentkus, V., and Zitikis, R. (1988): A remark on the Cramér–von Mises–Smirnov test. Liet. Mat. Rink. **28**, No. 1, 14–22. English transl.: Lith. Math. J. **28**, No. 1, 8–13. Zbl. 647.62032

Bentkus, V., and Zitikis, R. (1990): Probabilities of large deviations for L-statistics. Liet. Mat. Rink. **30**, No. 3, 479–488. English transl.: Lith. Math. J. **30**, No. 3, 215–222. Zbl. 706.62015

Bentkus, V., Götze, F., and Zitikis, R. (1993): Asymptotic expansions in the integral and local limit theorems in Banach spaces with applications to co-statistics. J. Theoretical Probab. **6**, 4, 727–780. Zbl. 807.60007

Bergström, H. (1944): On the central limit theorem. Skand. Aktuarietidskrift **27**, 139–153. Zbl. 706.62015

Bergström, H. (1951): On asymptotic expansions of probability functions. Skand. Aktuarietidskrift **1**. No. 2, 1–34. Zbl. 045.07301

Bernotas, V. (1979): Uniform and nonuniform proximity bounds for the distribution of two normalized sums of independent random variables with values in Banach spaces. Liet. Mat. Rink. **19**, No. 4, 55–68. English transl.: Lith. Math. J. **19**, No. 4, (1980), 482–489. Zbl. 424.60007

Berry, A.C. (1941): The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Am. Math. Soc. **49**, 122–136. Zbl. 025.34603

Bézandry, P.H., and Fernique, X. (1992): Sur la proprieté de la limite centrale dans $D[0,1]$. Ann. Inst. Henri Poincaré, Probab. Stat. **28**, No. 1, 31–46. Zbl. 749.60003

Bhattacharya, R.N. (1977): Refinements of the multidimensional central limit theorem and its application. Ann. Probab. **5**, 1–27. Zbl. 361.60001

Bhattacharya, R.N., and Denker, M. (1990): Asymptotic Statistics, Birkhäuser, Basel-Boston-Berlin. Zbl. 706.62049

Bhattacharya, R.N., and Rango Rao, R. (1976): Normal Approximation and Asymptotic Expansions. John Wiley, New York. Zbl. 331.41023

Bickel, P.J., Götze, F., and van Zwet, W.R. (1986): The Edgeworth expansion for U-statistics of degree two. Ann. Stat. **14**, 1463–1484. Zbl. 614.62015

Bikelis, A. (1967): Remainder terms in asymptotic expansions for characteristic functions and their derivatives. Liet. Mat. Rink. **7**, No. 4, 571–582. English transl.: Selected Transl. Math. Stat. Probab. **11**, 149–162 (1973). Zbl. 167.17401

Billingsley, P. (1969): Convergence of Probability Measures. John Wiley, New York. Zbl. 172.21201

Bloznelis, M. (1989): A lower bound for the convergence rate in the central limit theorem in Hilbert space. Liet. Mat. Rink. **29**, No. 4, 674–681. English transl.: Lith. Math. J. **29**, No. 4, 333–338. Zbl. 705.60010

Bloznelis, M., and Paulauskas, V. (1994): A note on the central limit theorem for stochastically continuous processes. Stoch. Proc. Appl. **53**, 351–361. Zbl. 838.60018

Bogachev, V.I. (1988): Subspaces of differentiability of smooth measures on infinite-dimensional spaces. Dokl. Akad. Nauk SSSR **299**, No. 1, 18–22. English transl.: Sov. Math. Dokl. **37**, 304–308. Zbl. 721.46030

Borisov, I.S. (1983): Problem of accuracy of approximation in the central limit theorem for empirical measures. Sib. Mat. Zh. **24**, No. 6, 14–25. English transl.: Sib. Math. J. **24**, 833–843. Zbl. 541.60020

Borisov, I.S. (1985): A remark on the speed of convergence in the central limit theorem in Banach spaces. Sib. Mat. Zh. **26**, No. 2, 29–35. English transl.: Sib. Math. J. **26**, 180–185. Zbl. 569.60004

Borisov, I.S. (1989): Approximation for distributions of smooth functionals of sums of independent random variables in Banach spaces. In: Asymptotic Analysis of Distributions of Stochastic Processes, Novosibirsk, Tr. Inst. Mat. **13**, 7–40. Engl. transl.: Sib. Adv. Math. **1**, No. 1, 1–38 (1991). Zbl. 718.60005, 705.60007

Borovkov, A.A. (1984): On the convergence rate in the invariance principle for a Hilbert space. Teor. Veroyatn. Primen. **29**, No. 3, 532–535. English transl.: Theory Probab. Appl. **29**, No. 3, 550–553. Zbl. 568.60008

Borovkov, A.A., and Sakhanenko, A.I. (1980): On estimates of the rate of convergence in the invariance principle for Banach spaces. Teor. Veroyatn. Primen. **25**, No. 4, 734–744. English transl.: Theory Probab. Appl. **25**, No. 4, 721–731. Zbl. 454.60033

Borovskikh, Yu.V., and Račkauskas, A. (1979): Asymptotic analysis of distributions in Banach spaces. Liet. Mat. Rink. **19**, No. 4, 39–54. English transl.: Lith. Math. J. **19**, 472–481. Zbl. 441.60009

Butzer, P.L., Hahn, L., and Roeckerath, M.T. (1979): General theorems on "little-o" rates of convergence of two weighted sums of independent Hilbert-space-valued random variables with applications. J. Multivariate Anal. **9**, 487–510. Zbl. 428.60028

Cartan, H. (1967): Calcul des formes differentielles. Hermann, Paris. Zbl. 184.12701

Chebotarev, V.I. (1982): Estimates of the convergence rate in the local limit theorem for the square of the norm in ℓ_2. In: Limit Theorems of Probability and Related Questions. Tr. Inst. Mat **1**, 122–126. Engl. transl.: Transl. Ser. Math. Eng. (1984), 219–225. New York, Optimization Software. Zbl. 508.60013

Chibisov, D.M. (1964): Some theorems on the limiting behavior of empirical distribution functions. Tr. Mat. Inst. Steklova **71**, 104–112 (1964). Engl. transl.: Selected Transl. Math. Stat. Probab. **6**, 147–156 (1966). Zbl. 163.40602

Csörgő, M., Csörgő, S., Horvath, L., and Mason, D.M. (1986): Weighted empirical and quantile processes. Ann. Probab. **14**, 31–85. Zbl. 589.60029

Csörgő, M., and Horvath, L. (1988): On the distributions of L_p norms of weighted uniform empirical and quantile processes. Ann. Probab. **16**. 142–161. Zbl. 646.62015

Csörgő, S. (1976): On an asymptotic expansion for the von Mises ω^2-statistics. Acta Sci. Math. **38**, 45–67. Zbl. 333.62020

Csörgő, S., and Stachó, L. (1980): A step toward asymptotic expansions for the Cramér–von Mises statistic. In: Analytic Function Methods in Probability Theory. Colloq. Math. Soc. James Bolyai **21**, 53–65. Zbl. 416.62015

Daletskii, Yu.L., and Fomin, S.V. (1983): Measures and Differential Equations in Infinite-Dimensional Spaces. Nauka, Moscow. Engl. transl.: Kluwer, Dordrecht (1991). Zbl. 536.46031

Daniels, H.E. (1945): The statistical theory of the strength of bundles of threads. Proc. R. Soc. Ser. A **183**, 405–435. Zbl. 063.01035

Davydov, Yu., and Lifshits, M.A. (1984): The fiber method in some probability problems. Itogi Nauki Tekh. Ser. Teor. Veroyatn. Mat. Stat. Kibern. **22**, 61–158. English transl.: J. Sov. Math. **31** (1985), 2796–2858. Zbl. 566.60040

Dehling, H. (1983): Limit theorems for sums of weakly dependent Banach-space-valued random variables. Z. Wahrscheinlichkeitstheorie Verw. Geb. **63**, 393–432. Zbl. 509.60012

Dudley, R.M. (1978): Central limit theorem for empirical measures. Ann. Probab. **6**, 899–923. Zbl. 404.60016

Dudley, R.M. (1987): Universal Donsker classes and metric entropy. Ann. Probab. **15**, 1306–1326. Zbl. 631.60004

Esseen, C.G. (1942): On the Liapounoff limit of error in the theory of probability. Ark. Mat., Astro. Fysik. **28A**, 1–19. Zbl. 027.33902

Esseen, C.G. (1945): Fourier analysis of distribution functions. Acta Math. **77**, 1–125. Zbl. 060.28705

Esseen, C.G. (1968): On the concentration function of a sum of independent random variables. Z. Wahrscheinlichkeitstheorie Verw. Geb. **9**, 290–308. Zbl. 195.19303

Feller, W. (1971): An Introduction to Probability Theory and Its Applications, Vol. II, 2nd. edn., John Wiley, New York. Zbl. 219.60003

Fernique, X. (1971): Regularité de processus Gaussien. Invent. Math. **12**, 304–320. Zbl. 217.21104

Fernique, X. (1994): Les functions aléatoires cadlag, la compacité de leurs lois. Liet. Mat. Rink. **34**, 288–306. Zbl. 829.60031

Fisz, M. (1959): A central limit theorem for stochastic processes with independent increments. Stud. Math. **18**, 223–227. Zbl. 152.16404

Gaenssler, P., and Stute, W. (1979): Empirical processes: a survey of results for independent and identically distributed random variables. Ann. Probab. **7**, 193–242. Zbl. 402.60031

Giné, E. (1976): Bounds for the speed of convergence in the central limit theorem in $C(S)$. Z. Wahrscheinlichkeitstheorie Verw. Geb. **36**, 317–331. Zbl. 351.60029

Giné, E., and Zinn, J. (1984): Some limit theorems for empirical processes. Ann. Probab. **12**, 929–989. Zbl. 553.60037

Götze, F. (1979): Asymptotic expansions for bivariate von Mises functionals. Z. Wahrscheinlichkeitstheorie Verw. Geb. **50**, 333–355. Zbl. 415.60008

Götze, F. (1981): On Edgeworth expansions in Banach spaces. Ann. Probab. **9**, 852–859. Zbl. 473.60009

Götze, F. (1983): On the rate of convergence in the CLT under moment conditions. Preprints in Statistics, Univ. Cologne.

Götze, F. (1984): Expansions for von Mises functionals. Z. Wahrscheinlichkeitstheorie Verw. Geb. **65**, 599–625. Zbl. 531.60037

Götze, F. (1985): Asymptotic expansion in a functional limit theorem. J. Multivariate Anal. **16**, 1–20. Zbl. 557.60027

Götze, F. (1986): On the rate of convergence in the central limit theorem in Banach spaces. Ann. Probab. **14**, 922–942. Zbl. 599.60009

Götze, F. (1987): Approximations for multivariate U-statistics. J. Multivariate Anal. **22**, 212–229. Zbl. 624.62028

Götze, F. (1989): Edgeworth expansions in functional limit theorems. Ann. Probab. **17**, 1602–1634. Zbl. 689.60038

Götze, F., and Hipp, C. (1982): Asymptotic expansions in the CLT under moment conditions. Z. Wahrscheinlichkeitstheorie Verw. Geb. **42**, 67–87. Zbl. 369.60027

Hahn, M.G. (1978): Central limit theorem in $D[0,1]$. Z. Wahrscheinlichkeitstheorie Verw. Geb. **44**, 89–101. Zbl. 364.60002

Hardy, G.H., and Littlewood, J.E. (1920): A new solution of Waring's problem. Quarterly J. Math. **48**, 272–293. Zbl. 47.0114.01

Helmers, R. (1982): Edgeworth expansions for linear combinations of order statistics. Amsterdam Math. Center Tracts **105**. Zbl. 485.62017

Ibragimov, I.A., and Linnik, Yu. (1971): Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff Publishing, Groningen. Zbl. 154.42201

Juknevičienė, D. (1985): Central limit theorem in the space $D[0,1]$. Liet. Mat. Rink. **25**, No. 3, 198–205. English transl.: Lith. Math. J. **25**, No. 3, 293–298. Zbl. 593.60032

Kandelaki, N.P. (1965): On a limit theorem in Hilbert space. Tr. Vychisl. Tsentra Akad. Nauk Gruz. SSR **5**, 46–55 (in Russian). Zbl. 253.60014

Kandelaki, N.P., and Vakhaniya, N.N. (1969): On a bound for the convergence rate in the central limit theorem in Hilbert space. Tr. Vychisl. Tsentra Akad. Nauk Gruz. SSR **9**, 150–160 (in Russian). Zbl. 198.22901

Kiefer, J. (1972): Skorokhod embedding of multivariate r.v.'s and the sample d.f. Z. Wahrscheinlichkeitstheorie Verw. Geb. **24**, 1–35. Zbl. 267.60034

Koroliuk, V.S., and Borovskikh, Yu.V. (1984): Asymptotic Analysis of Distributions of Statistics. Naukova Dumka, Kiev (in Russian). Zbl. 565.62003

Kuelbs, J., and Kurtz, T. (1974): Berry-Esseen estimates in Hilbert space and an application to the law of the iterated logarithm. Ann. Probab. **2**, 387–407. Zbl. 298.60017

Kukuš, A.G. (1981): The weak convergence of measures and the convergence of cumulants. Teor. Veroyatn. Mat. Stat. **23**, 74–80 (1980). English transl.: Theory Probab. Math. Stat. **25**, 79–86. Zbl. 482.28010

Kukuš, A.G. (1982): The central limit theorem in Hilbert space in terms of the Lévy–Prokhorov metric. Teor. Veroyatn. Mat. Stat. **25**, 55–63. English transl.: Theory Probab. Math. Stat. **25**, 61–68. Zbl. 457.60005

Lapinskas, R. (1978): Approximation of partial sums in certain Banach spaces. Liet. Mat. Rink. **18**, No. 4, 65–71. English transl.: Lith. Math. J. **18**, No. 4, 494–498. Zbl. 403.60016

Liapunov, A.M. (1900): Sur une proposition de la théorie des probabilités. Bull. Acad. Sci., St. Petersbourg (5) **13**, 359–386. Jbuch 31.0228.02

Liapunov, A.M. (1901): Nouvelle forme du théoreme sur la limite de théorie des probabilités. Mem. Acad. Sci. St. Petersbourg (8) **12**, 1–24. Jbuch 33.0248.07

Lifshits, M.A. (1983): Distribution (density) of the maximum of a Gaussian process. Teor. Veroyatn. Primen. **31**, No. 1, 131–142. English transl.: Theory Probab. Appl. **31**, No. 1, 125–132. Zbl. 602.60040

Lindeberg, J.W. (1920): Über das Exponentialgesetz in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fenn., Ser. A. **16**, 1–23. Jbuch 47.0485.01

Lindeberg, J.W. (1922): Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. Math. Z. **15**, 211–225. Jbuch 48.0602.04

Liubinskas, K. (1987): On the closeness of moments in the central limit theorem in Banach spaces. Liet. Mat. Rink. **27**, No. 2, 285–302. Zbl. 635.60006

Marcus, M.B., and Pisier, G. (1981): Random Fourier Series with Applications to Harmonic Analysis. Ann. Math. Stud. **101**. Zbl. 474.43004

Marcus, M.B., and Shepp, L.A. (1972): Sample behavior of Gaussian processes. Proc. Sixth Berkeley Symp. Math. Stat. Probab. **3**, 423–442. Zbl. 379.60040

Martynov, G.V. (1978): Omega-Square Tests, Nauka, Moscow (In Russian).

Massart, P. (1986): Rates of convergence in the central limit theorem for empirical processes. Ann. Inst. Henri Poincaré, Probab. Stat. **22**, 381–423. Zbl. 615.60032

Massart. P. (1989): Strong approximation for multivariate empirical and related processes, via KMT construction. Ann. Probab. **17**, 266–291. Zbl. 675.60026

Nagaev, S.V. (1976): An estimate of the remainder term in the multi-dimensional central limit theorem. Proc. Third Japan-USSR Symp. Probab. Theory; Lect. Notes Math. **550**, 419–438. Zbl. 363.60024

Nagaev, S.V. (1983): On accuracy of normal approximation for the distribution of a sum of independent Hilbert-space-valued random variables. Lect. Notes Math. **1021**, 461–473. Zbl. 526.60008

Nagaev, S.V. (1985): On the rate of convergence to the normal law in Hilbert space. Teor. Veroyatn. Primen. **30**, No. 1, 19–32. English transl.: Theory Probab. Appl. **30**, No. 1, 19–37 (1986). Zbl. 657.60009

Nagaev, S.V. (1989a): A Berry-Esseen-type estimate for sums of random variables with values in a Hilbert space. Sib. Mat. Zh. **30**, No. 3, 84–96. English transl.: Sib. Math. J. **30**, No. 3, 413–423. Zbl. 675.60010

Nagaev, S.V. (1989b): A new approach to the study of the distribution of the norm of a random element in Hilbert space. Abst., Fifth Vilnius Conf. Probab.Th. Math. Stat. 77–78. English transl.: Vol. II, 214–226 (1990). Zbl. 733.60013

Nagaev, S.V., and Chebotarev, V.I. (1978): Bounds for the convergence rate in the central limit theorem in the Hilbert space ℓ_2. In: Mathematical Analysis and Related Topics, Novosibirsk, 153–182 (In Russian).

Nagaev, S.V., and Chebotarev, V.I. (1986): A refinement of the error estimate of the normal approximation in a Hilbert space. Sib. Mat. Zh. **27**, No. 3, 154–173. English transl.: Sib. Math. J. **27**, No. 3, 434–450. Zbl. 604.60009

Nagaev, S.V., and Chebotarev, V.I. (1987): Asymptotic expansions for the distribution of the sum of i.i.d. Hilbert-space-valued r.v.'s. Probab. Theory Math. Stat. (1985), Vilnius. English transl.: VNU Science Press, Utrecht, The Netherlands **2**, 357–363. Zbl. 673.60007

Nagaev, S.V., and Chebotarev, V.I. (1989a): On asymptotic expansions of Bergström type in Hilbert space. In: Asymptotic Analysis of Distributions of Stochastic Processes. Tr. Inst. Mat. **13**, 66–77 (1989). English transl.: Sib. Adv. Math. **1**, No. 2, 130–145 (1991). Zbl. 705.60008

Nagaev, S.V., and Chebotarev, V.I. (1989b): On Edgeworth expansions in Hilbert space. Tr. Inst. Mat. **20**, 170–203 (1993). English transl.: Sib. Adv. Math. **3**, No. 3, 89–122 (1993). Zbl. 866.60012

Nikitin, Ya.Yu. (1972): On a boundary problem for an empirical process. Dokl. Akad. Nauk SSSR **205**, No. 5, 1043–1045. English transl.: Sov. Math. Dokl. **13**, No. 4, 1081–1084. Zbl. 306.60014

Norvaiša, R. (1993): The central limit theorem for empirical and quantile processes in some Banach spaces (Preprint). English transl.: Stochastic Processes Appl. **46**, No. 1, 1–27 (1993). Zbl. 780.60011

Norvaiša, R., and Paulauskas, V. (1991): Rate of convergence in the central limit theorem for empirical processes. J. Theor. Probab. **4**, No. 3, 511–534. Zbl. 734.60020

Norvaiša, R., and Zitikis, R. (1991): Asymptotic behavior of linear combinations of functions of order statistics. J. Stat. Plann. Inference, 305–317. Zbl. 62046

O'Reilly, N. (1974): On the convergence of empirical processes in supernorm metrics. Ann. Probab. **2**, 642–651. Zbl. 60007

Orlov, A.I. (1974): The rate of convergence of the distribution of the Smirnov-von Mises statistic. Teor. Veroyatn. Primen. **19**, No. 4, 766–786. English transl.: Theory Probab. Appl. **19**, No. 4, 737–757. Zbl. 301.60007

Osipov, L.V. (1977): On large deviations of sums of independent random vectors. Abst., 2nd Vilnius Conf. Probab. Th. Math. Stat. **2**, 95–96 (in Russian).

Osipov, L.V. (1978a): On large deviation probabilities for sums of independent random vectors. Teor. Veroyatn. Primen. **23**, No. 3, 510–526. English transl.: Theory Probab. Appl. **23**, No. 3, 490–506 (1979). Zbl. 437.60005

Osipov, L.V. (1978b): Large Deviation Probabilities for Sums of Independent Random Vectors. Doct. Dissertation, Leningrad (in Russian).

Osipov, L.V., and Rotar', V.I. (1984): On an infinite-dimensional central limit theorem. Teor. Veroyatn. Primen. **29**, No. 2, 366–373. English transl.: Theory Probab. Appl. **29**, No. 2, 375–382. Zbl. 544.60015

Paulauskas, V. (1973): On the concentration function of finite-dimensional and infinite-dimensional random vectors. Liet. Mat. Rink. **13**, No. 1, 137–157. English transl.: Lith. Math. J. **13**, No. 1, 97–111. Zbl. 264.60033

Paulauskas, V. (1976a): Estimate of the convergence rate in the central limit theorem in $C(S)$. Liet. Mat. Rink. **16**, No. 4, 167–201. English transl.: Lith. Math. J. **16**, No. 4, 587–611. Zbl. 392.60022

Paulauskas, V. (1976b): On the rate of convergence in the central limit theorem in certain Banach spaces. Teor. Veroyatn. Primen. **21**, No. 4, 775–791. English transl.: Theory Probab. Appl. **21**, No. 4, 754–769. Zbl. 403.60007

Paulauskas, V. (1981): Estimation of the convergence rate in the central limit theorem in ℓ_p. Liet. Mat. Rink. **21**, No. 1, 109–119. English transl.: Lith. Math. J. **21**, No. 1, 55–62. Zbl. 455.60031

Paulauskas, V. (1984): On the central limit theorem in Banach space c_0. Probab. Math. Stat. **3**, 127–141.

Paulauskas, V. (1990): On the rate of convergence for the weighted empirical process. Probability in Banach Spaces **7**. Proc. 7th Int. Conf. Oberwolfach, FRG 1988, Prog. Probab. **21**, 147–158 (1990). Zbl. 704.60009

Paulauskas, V., and Juknevičienė, D. (1988): On the rate of convergence in the central limit theorem in $D[0, 1]$. Liet. Mat. Rink. **28**, No. 3, 507–519. English transl.: Lith. Math. J. **28**, No. 3, 229–239. Zbl. 711.60008, Zbl. 657.60011

Paulauskas, V., and Račkauskas, A. (1987): Approximation Theory in the Central Limit Theorem. Exact Results in Banach Spaces, Mokslas, Vilnius. English transl.: Kluwer, Dordrecht (1989). Zbl. 708.60005

Paulauskas, V., and Račkauskas, A. (1991): Nonuniform estimates in the central limit theorem in Banach spaces. Liet. Mat. Rink. **31**, No. 3, 483–496. Zbl. 786.60009

Paulauskas, V., and Stieve, Ch. (1990): On the central limit in D[0, 1] and D([0, 1], H). Liet. Mat. Rink. **30**, No. 3, 567–579. English transl.: Lith. Math. J. **30**, No. 3, 267–276. Zbl. 722.60023

Petrov, V.V. (1975): Sums of Independent Random Variables, Nauka, Moscow (1972). English transl.: Springer-Verlag, Berlin Heidelberg New York

Phoenix, S.L., and Taylor, H. (1973): The asymptotic strength distribution of a general fiber bundle. Adv. Appl. Probab. **5**, 200–216. Zbl. 272.60006

Prokhorov, Yu.V., and Sazonov, V.V. (1969): On estimates of the convergence rate in the central limit theorem in the infinite-dimensional case. In: Soviet-Japanese Khabarovsk Symposium on Probability Theory **1**, 223–230.

Puri, M.L., and Seoh, M. (1987): On the rate of convergence in normal approximation and large deviation probabilities for a class of statistics. Teor. Veroyatn. Primen. **33**, No. 4, 736–750. Already in English: Theory Probab. Appl. **33**, No. 4, 682–697. Zbl. 665.62017

Rachev, S.T., and Rüschendorf, L. (1992): Rate of convergence for sums and maxima and doubly ideal metrics. Teor. Veroyatn. Primen. **37**, No. 2. Already in English: Theory Probab. Appl. **37**, No. 2, 222–235. Zbl. 787.60006

Rachev, S.T, and Yukich, J.E. (1989): Rates for the CLT via new ideal metrics. Ann. Probab **17**, 775–788. Zbl. 675.60018

Račkauskas, A. (1981): Approximation in the uniform metric of sums of independent random variables with values in Hilbert space. Liet. Mat. Rink. **21**, No. 3, 83–90. English transl.: Lith. Math. J. **21**, No. 3, 258–263. Zbl. 497.60008

Račkauskas, A. (1988): Probabilities of large deviations in Linnik zones in Hilbert space. Liet. Mat. Rink. **28**, No. 3, 520–533. English transl.: Lith. Math J. **28**, No. 3, 239–248. Zbl. 657.60008

Račkauskas, A. (1991): On the convergence rate in martingale CLT in Hilbert space. Liet. Mat. Rink. **31**, No. 3, 497–512. English transl.: Lith. Math. J. **31**, No. 3, 345–355. Zbl. 777.60027

Rhee, W.S., and Talagrand, M. (1984): Bad rates of convergence for the CLT in Hilbert space. Ann. Probab. **12**, 843–850. Zbl. 545.60014

Rhee, W.S., and Talagrand, M. (1986): Uniform convexity and the distribution of the norm of a Gaussian measure. Probab. Theory Relat. Fields **71**, 59–68. Zbl. 554.60007

Rosenkrantz, W.A. (1969): A rate of convergence for the von Mises statistic. Trans. Am. Math. Soc. **139**, 329–337. Zbl. 182.52301

Sakalauskas, V. (1983): Approximation by a stable law in nonuniform metrics of Lévy–Prokhorov and χ type. Liet. Mat. Rink. **23**, No. 4, 40–49. English transl.: Lith. Math. J. **23**, No. 4, 384–391. Zbl. 561.60013

Sakhanenko, A.I. (1988): Simple method of obtaining estimates in the invariance principle. Probab. Theory Math. Stat. (Kyoto 1986). Lect. Notes Math. **1299**, 430–443. Zbl. 637.60010

Sazonov, V.V. (1968): On ω^2-test. Sankhya, Ser. A. **30**, 205–209. Zbl. 177.47201

Sazonov, V.V. (1969): An improvement of a convergence-rate estimate. Teor. Veroyatn. Primen. **14**, No. 4, 667–678. English transl.: Theory Probab. Appl. **14**, No. 4, 640–651. Zbl. 185.46402

Sazonov, V.V. (1981): Normal Approximation – Some Recent Advances. Lect. Notes Math. **879**. Springer-Verlag. Zbl. 462.60006

Sazonov, V.V and Ul'yanov, V.V. (1990): Speed of convergence in the central limit theorm in Hilbert space under weakened moment conditions. In: Probab. Theory Math. Stat.; Proc. Fifth Vilnius Conf. **2**, 394–410. Vilnius, Mokslas. Zbl. 734.60008

Sazonov, V.V., and Ul'yanov, V.V. (1991): An improved estimate of the accuracy of the Gaussian approximation in Hilbert space. In: New Trends in Probab., and Stat., Mokslas, VSP, 123–136. Zbl. 803.60006

Sazonov, V.V., Ul'yanov, V.V., and Zalesskii, B.A. (1987a): On normal approximation in Hilbert space. In: Probab. Theory Math. Stat. (Vilnius 1985). Proc. 4th Vilnius Conf., VNU Science Press, Utrecht, The Netherlands, 561–580. Zbl. 652.60011

Sazonov, V.V., Ul'yanov, V.V., and Zalesskii, B.A. (1987b): Asymptotic expansions refining the central limit theorem in Hilbert space. In: Probab. Theory Appl. (Yu.A. Prokhorov and V.V. Sazonov, Eds.), VNU Science Press, Utrecht, The Netherlands **1**, 679–688. Zbl. 697.60007

Sazonov, V.V., Ul'yanov, V.V., and Zalesskii, B.A. (1988): Normal approximation in Hilbert space, I, II. Teor. Veroyatn. Primen. **33**, No. 2, 225–245; No. 4, 733–754. English transl.: Theory Probab. Appl. **33**, No. 2, 207–227; No. 4, 473–487, No. 2. Zbl. 649.60005, No. 4. Zbl. 662.60009

Sazonov, V.V., Ul'yanov, V.V., and Zalesskii, B.A. (1989a): Asymptotically precise estimate of the accuracy of Gaussian approximation in Hilbert space. J. Multivariate Anal. **28**, 304–330. Zbl. 675.60011

Sazonov, V.V., Ul'yanov, V.V., and Zalesskii, B.A. (1989b): A precise estimate of the convergence rate in the central limit theorem in Hilbert space. Mat. Sb. **180**, No. 12, 1587–1613. English transl.: Math USSR Sb. **68** (1991), No. 2, 453–482. Zbl. 709.60006

Sazonov, V.V., and Zalesskii, B.A. (1985): On the CLT in Hilbert space. J. Multivariate Anal. **24**, 495–526. Zbl. 603.60018

Schmidt, W. (1984): Bounds for exponential sums. Acta Arith. **44**, 281–297. Zbl. 544.10036

Senatov, V.V. (1981): Some lower convergence rate bounds in the central limit theorem. Dokl. Akad. Nauk SSSR **256**, No. 6, 1318–1321. English transl.: Sov. Math. Dokl. **23**, 188–192. Zbl. 603.60018

Senatov, V.V. (1983): On estimating the rate of convergence in the central limit theorem over a system of balls in R^k. Teor. Veroyatn. Primen. **28**, No. 2, 440–445. English transl.: Theory Probab. Appl. **28**, No. 2, 463–467. Zbl. 569.60022

Senatov, V.V. (1985a): On the dependency of estimates of the convergence rate in the central limit theorem on the covariance operator of the summands. Teor. Veroyatn. Primen. **30**, No. 2, 354–357. English transl.: Theory Probab. Appl. **30**, No. 2, 380–383. Zbl. 569.60023

Senatov, V.V. (1985b): Four examples of lower estiimates in the multi-dimensional central limit theorem. Teor. Veroyatn. Primen. **30**, No. 4, 750–758. English transl.: Theory Probab. Appl. **30**, No. 4, 797–805. Zbl. 579.60020

Senatov, V.V. (1986): On the dependence of estimates of the convergence rate in the central limit theorem for balls with center zero on the covariance operator of the summands. Teor. Veroyatn. Primen. **31**, No. 1, 128–132. English transl.: Theory Probab. Appl. **31**, No. 1, 119–122. Zbl. 589.60004

Senatov, V.V. (1989a): On bounds for the convergence rate in the central limit theorem in Hilbert space. Abst., Fifth Vilnius Conf. Theory Probab. Math. Stat. **4**. (In Russian)

Senatov, V.V. (1989b): On estimating the rate of convergence in the central limit theorem in Hilbert space. Lect. Notes Math. **1412**, 309–327. Zbl. 692.60005

Serfling, R.J. (1980): Approximation Theorems of Mathematical Statistics, John Wiley, New York.

Siegel, G. (1981): Upper estimates for the concentration function in Hilbert space. Teor. Veroyatn. Primen. **26**, No. 2, 335–349. English transl. Theory Probab. Appl. **26**, No. 2, 328–343. Zbl. 487.60014

Smirnov, N.V. (1937): On the distribution of the ω^2-test of von Mises. Rec. Math. (NS) **2**, 973–993 (In Russian). Zbl. 018.41202

Stigler, S.M. (1974): Linear functions of order statistics with smooth weight functions. Ann. Stat **2**, 676–693. Zbl. 286.62028

Sweeting, T.J. (1977): Speed of convergence for the multi-dimensional central limit theorem. Ann. Probab. **5**, 28–41. Zbl. 362.60041

Sweeting, T.J. (1980): Speeds of convergence and asymptotic expansions in the CLT – a treatment by operators. Ann. Probab. **8**, 279–281. Zbl. 444.60017

Thomasian, A. (1969): The Structure of Probability Theory with Applications, McGraw-Hill, New York. Zbl. 204.50101

Trotter, H.F. (1959): Elementary proof of the central limit theorem. Arch. Math. **10**, 226–234. Zbl. 086.34002

Ul'yanov, V.V. (1981): An estimate for the rate of convergence in the central limit theorem in a separable Hilbert space. Mat. Zametki **29**, No. 1, 145–153. English transl.: Math. Notes **29**, No. 1, 78–82. Zbl. 458.60008

Ul'yanov, V.V. (1986): Asymptotic expansions for distributions of sums of independent random variables in H. Teor. Veroyatn. Primen. **31**, No. 1, 31–46. English transl.: Theory Probab. Appl. **31**, No. 1, 25–39. Zbl. 625.60045

Vakhaniya, N.N., Tarieladze, V.I., and Chobanyan, S.A. (1987): Probability Distributions in Banach Spaces, Nauka, Moscow (1985). English transl.: D. Riedel Publishing Company, Dordrecht. Zbl. 572.60003

Vandemaele, M., and Veraverbeke, N. (1982): Cramér-type large deviations for linear combinations of order statistics. Ann. Probab. **10**, 423–434. Zbl. 482.60026

Vinogradov, I.M. (1934): A new evaluation of $G(n)$ in Waring's problem. Dokl. Akad. Nauk SSSR **5**, 249–251. English transl.: Version 251–253. Zbl. 011.00803

Vinogradova, T.R. (1985): On the accuracy of normal approximation on sets defined by a smooth function, I, II. Teor. Veroyatn. Primen. **30**, No. 2, 219–229; **30**, No. 3, 554–557. English transl.: Theory Probab. Appl. **30**, No. 2, 235–246; **30**, No. 3, 590–593. No. 2 Zbl. 573.60010, No. 3 Zbl. 658.60048

Wenocur, R.S., and Dudley, R.M. (1981): Some special Vapnik-Chervonenkis classes. Discrete Math. **33**, 313–318. Zbl. 459.60008

Weyl, H. (1916): Über die Gleichverteilung der Zahlen mod-Eins, Ann. **77**, 313–352. Jbuch 46.0278.06

Yurinskii, V.V. (1977): On the error in Gaussian approximation of convolutions. Teor. Veroyatn. Primen. **22**, No. 2, 242–253. English transl.: Theory Probab. Appl. **22**, No. 2, 236–247. Zbl. 378.60008

Yurinskii, V.V. (1981): An error estimate for the normal approximation of the probability of landing in a ball. Dokl. Akad. Nauk SSSR **258**, No. 3, 557–558. English transl.: Sov. Math. Dokl. **23**, 576–578. Zbl. 508.60026

Yurinskii, V.V. (1982): On the accuracy of normal approximation of the probability of hitting a ball. Teor. Veroyatn. Primen. **27**, No. 2, 270–278. English transl.: Theory Probab. Appl. **27**, No. 2, 280–289. Zbl. 565.60005

Yurinskii, V.V. (1983): Error of normal approximation. Sib. Mat. Zh. **24**, No. 6, 188–199. English transl.: Sib. Math. J. **24**, No. 6, 977–987. Zbl. 541.60019

Yurinskii, V.V. (1991): On asymptotic analysis of large deviations in Hilbert space, I, II. Teor. Veroyatn. Primen. **36**, No. 1, 78–92; No. 3, 535–541. English transl.: Theory Probab. Appl. **36**, No. 1, 99–114; No. 3, 548–554. No. 1 Zbl. 727.60026, No. 3 Zbl. 813.60007

Zaĭtsev, A.Yu. (1987): On the Gaussian approximation of convolutions under multidimensional analogues of S.N. Bernstein's inequality conditions. Probab. Theory Relat. Fields **74**, 535–566. Zbl. 612.60031

Zalesskii, B.A. (1982): Estimation of the accuracy of normal approximation in Hilbert space. Teor. Veroyatn. Primen. **27**, No. 2, 279–285. English transl.: Theory Probab. Appl. **27**, No. 2, 290–298. Zbl. 565.60004

Zalesskii, B.A. (1985): On the convergence rate in the central limit theorem on a class of sets in Hilbert space. Teor. Veroyatn. Primen. **30**, No. 4, 662–670. English transl.: Theory Probab. Appl. **30**, No. 4, 702–711. Zbl. 586.60010

Zalesskii, B.A. (1988): On the accuracy of normal approximation in Banach spaces. Teor. Veroyatn. Primen. **33**, No. 2, 257–265. English transl.: Theory Probab. Appl. **33**, No. 2, 239–247. Zbl. 666.60009

Zalesskii, B.A. (1989): Probabilities of large deviations in Hilbert space. Teor. Veroyatn. Primen. **34**, No. 4, 650–655. English transl.: Theory Probab. Appl. **34**, No. 4, 591–596. Zbl. 695.60029

Zalesskii, B.A., and Sazonov, V.V. (1984): Closeness of moments for normal approximation in a Hilbert space. Teor. Veroyatn. Primen. **28**, No. 2, 251–263. English transl.: Theory Probab. Appl. **28**, No. 2, 263–277. Zbl. 515.60013

Zalesskii, B.A., Sazonov, V.V., and Ul'yanov, V.V. (1988): An asymptotically regular estimate of the accuracy of normal approximation in Hilbert space. Teor. Veroyatn. Primen. **33**, No. 4, 753–754. English transl.: Theory Probab. Appl. **33**, No. 4, 700–701. Zbl. 662.60009

Zitikis, R. (1988): Asymptotic expansions in the local limit theorem for ω_n^2 statistics. Liet. Mat. Rink. **28**, No. 3, 461–474. Zbl. 662.62052

Zitikis, R. (1989): Asymptotic expansions for the derivatives of the distribution function of the Anderson-Darling statistic. Liet. Mat. Rink. **29**, No. 1, 35–53. Zbl. 777.62026

Zitikis, R. (1990a): Smoothness of the distribution function of the FL-statistic I, II. Liet. Math. Rink. **30**, No. 2, 233–246; **30**, No. 3, 500–512. English transl.: Lith. Math. J. **30**, No. 2, 97–106; **30**, No. 3, 231–239. No. 2 Zbl. 716.62028, No. 3. Zbl. 716.62029

Zitikis, R. (1990b): A uniform limit theorem for the densities of L-statistics. Liet. Mat. Rink **30**, No. 4, 728–740. English transl.: Lith. Math. J. **30**, No. 4, 331–341. Zbl. 761.62058

Zitikis, R. (1991a): On large deviations for L-statiistics. NewTrends in Probability and Statistics **1**, (V.V. Sazonov and T. Shervashidze, eds.), VSP Mokslas, Vilnius, 137–164. Zbl. 768.60021

Zitikis, R. (1991b): Cramér-type large deviations for a class of statistics. Liet. Mat. Rink. **31**, No. 2, 302–310. English transl.: Lith. Math. J. **31**, No. 2, 204–210. Zbl. 738.62028

Zolotarev, V.M. (1976a): Metric distances in spaces of random variables and their distributions. Mat. Sb. **101**, No. 3, 416–450. English transl.: Math. USSR, Sb. **101**, No. 3, 373–401. Zbl. 376.60003

Zolotarev, V.M. (1976b): Approximations of distributions of sums of independent random variables with values in infinite-dimensional spaces. Teor. Veroyatn. Primen. **21**, No. 4, 741–758. English transl.: Theory Probab. Appl. **21**, No. 4, 721–737. Zbl. 378.60003

Zolotarev, V.M. (1977): Ideal metrics in the problem of approximating distributions of sums of independent random variables. Teor. Veroyatn. Primen. **22**, No. 3, 449–465. English transl.: Theory Probab. Appl. **22**, No. 3, 433–449. Zbl. 385.60025